Aalto University
School of Business

The Market Value of ESG during COVID-19 Market Crash in Europe

Master's Thesis
Eetu Rasimäki
Aalto University School of Business
Master's Programme in Finance
2021

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author	Eetu Rasimäki
Title of thesis	The Market Value of ESG during COVID-19 Market Crash in Europe
Degree	Master of Science in Economics and Business Administration
Degree programme	Master's Programme in Finance
Thesis advisor(s)	Markku Kaustia
Year of approval	2021
Number of pages	62
Language	English

Abstract

This thesis studies the impact of firm's environmental, social and governance policies on stock returns during the 2020 market crash, which was caused by the exogenous shock of global pandemic that emerged during the first quarter of 2020. Past research has already provided theories and empirical evidence for stock market overperformance, during market crisis, for firms that are observed to be environmentally and socially more responsible. This thesis shows that European firms with higher Social performance has higher returns during the COVID-19 market crash, even after controlling for various factors. On the other hand, good Governance practices of the firm predicts negative impact on returns during the crash. Firm's environmental performance has no significant impact on returns. Also, when controlling for the location of the firm based on the COVID-19 situation in the country where the firm is headquartered, I observe these same effects only for firms located in countries with worse COVID-19 situation. In addition, firms with high Social performance and which are located in countries with worse COVID-19 situation experienced underperformance during the recovery period after the market crash and after overperforming during the crash period. These findings suggest that investor could make his portfolio more resilient to large market crashes by buying stocks of firms with high Social performance, but he might have to pay for this by poorer returns after the crash.

Keywords ESG, CSR, COVID-19, market crash

Aalto-yliopisto
Kauppakorkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Maisterintutkinnon tutkielman tiivistelmä

Tekijä Eetu Rasimäki

Työn nimi ESG:n markkina-arvo COVID-19 markkinaromahduksen aikana Euroopassa

Tutkinto Kauppatieteiden maisteri

Koulutusohjelma Rahoituksen maisteriohjelma

Työn ohjaaja(t) Markku Kaustia

Hyväksymisvuosi 2021 **Sivumäärä** 62 **Kieli** Englanti

Tiivistelmä

Tämä tutkielma tutkii yrityksen ympäristöllisen, sosiaalisen ja hallinnolisten käytäntöjen vaikutusta osaketuottoihin 2020 markkinaromahduksen aikana, joka oli seurausta globaalin pandemian aiheuttamasta eksogeenisesta shokista 2020 ensimmäisen kvartaalin aikana. Aikaisemmat tutkimukset ovat tarjonneet teorioita ja empiiristä näyttöä ympäristöllisesti ja sosiaalisesti vastuullisten yritysten ylituotosta osakemarkkinoilla markkinaromahdusten aikana. Tämä tutkielma osoittaa, että Eurooppalaiset yritykset, jotka ovat sosiaalisesti vastuullisia tuottivat muita keskimääräisesti enemmän COVID-19 markkinaromahduksen aikana. Toisaalta yritykset jotka ovat hallinnollisesti muita parempia tuottivat keskimäärin vähemmän kyseisen markkinaromahduksen aikana. Yrityksen ympäristöllisellä vastuullisuudella ei ollut merkittävää vaikutusta tuottoihin. Kontroloitaessa yrityksen sijainnin perustuen, jakaen yritykset sen perusteella kuinka vakava COVID-19 tilanne kunkin yrityksen pääkonttorin maassa on, löydän samat vaikutukset vain yrityksille, jotka sijaitsevat maissa joissa on huonompi COVID-19 tilanne. Lisäksi, yritykset jotka toimivat sosiaalisesti vastuullisesti ja jotka sijaitsevat maissa joissa on huonompi COVID-19 tilanne kokivat keskimääräistä huonompia tuottoja markkinaromahduksen jälkeisen elpymisjakson aikana ja sen jälkeen kun kyseiset yritykset olivat pärjänneet keskimääräistä paremmin markkinaromhaduksen aikana. Nämä löydökset vihjaa, että sijoittaja voi tehdä sijoitussalkustaan kestävämmän markkinaromahdusten suhteen, ostamalla sosiaalisesti vastuullisia yrityksiä, mutta hän saattaisi joutua maksamaan tästä huonommalla tuotto-odotuksella romahduksen jälkeisenä ajanjaksona.

Avainsanat ESG, CSR, COVID-19, markkinaromahdus

Contents

Tables & Figures

1 Introduction

"The social responsibility of business is to increase its profits". This is the title of Milton Friedman's article published in 1970. Friedman's view of businesses as solely shareholder wealth maximizers has experienced push back, especially in recent years, by more wholistic view of stakeholder capitalism that serves the interests of all the stakeholders of the company. In addition to the well-being of the shareholders, also the interests of customers, employees and other stakeholders should be taken into account in firm's decision-making process. One way to justify this new approach to business is to provide evidence that social responsibility increases profitability and stock returns instead of creating agency problems between shareholders and managers as Friedman (1970) suggests.[1] The current prevailing view of social responsibility among executives and investment professionals is that environmental, social and governance (ESG) policies maximize shareholder wealth (see McKinsey & Company (2020)). This view has been supported also by academic literature.[2]

The COVID-19 pandemic and the related stock market crash of early 2020 provide unique opportunity to study the return behavior of stocks during the market crash, and test how different factors and firm characteristics might impact stock returns. As Albuquerque et al. (2020) mention, companies did not have much time to respond to the COVID-19 crisis and thus it is possible to study the causality from firm characteristics to stock returns. The COVID-19 crisis was an unexpected exogenous shock that also resulted into a stock market crash. Therefore, the stock market crash that followed the realized and expected effects of the pandemic makes it possible to study the relationship between ESG policies and stock returns during a market crash period, and the exogenous nature of the shock also warrants the study of causality from ESG policies to returns.

[1] There is also view proposed by Hart and Zingales (2017) that suggest that the focus of the firm should be on shareholder welfare and not shareholder wealth maximization. The argument suggests that the shareholders have other goals besides value maximization, and therefore companies could let the shareholders vote on broad outlines of corporate policy. If this is the case it would not be necessary to justify better ESG policies by showing them to be positively related to higher market value.

[2] See for example Friede et al. (2015) and Margolis et al. (2009) for meta-studies of the relationship between ESG and corporate financial performance.

Past research has suggested that stocks of firms that operate in environmentally and socially responsible way have been more resilient during market shocks. Therefore, owning stocks of these firms can provide partial hedge during systematic market shocks. Lins et al. (2016) provide evidence for this kind of outperformance by firms with high environmental and social performance during the 2008-2009 financial crisis. The social, environmental and governance performance of a firm can be proxied by scores provided by one of the multiple rating agencies. The goal of these rating agencies is to approximate the ESG performance of an individual firm by using various metrics related to the activities of the firm. Usually, the rating agency provides three high-level scores (pillar scores) of Environmental, Social and Governance performance. These three scores can be aggregated to create the overall ESG score. In this thesis I focus on analysing these three scores individually. By not aggregating them to one score, as is often done, I can study the effect of each pillar individually. I use ESG scores provided by Refinitiv (Thomson Reuter).

There is already evidence that socially responsible firms outperformed other firms during the COVID-19 market crash (see Albuquerque et al. (2020) and Jurvanen (2020)). My goal is to test whether the three ESG pillars had any impact on the returns during the COVID-19 market crash in European markets.

Multiple models have been suggested to explain the positive relationship between social responsibility and stock returns during market crisis. Lins et al. (2017) suggest that firm can create social capital by environmental and social (ES) activities and higher social capital makes the stock returns of the firm more resilient during market crisis. Albuquerque et al. (2019) develop a model which shows that ES activities increase market differentation and thus creates more loyal customer base that benefit the firm's financial performance during poor economic times and decreased consumer demand. Third hypothesis that tries to explain the outperformance of high ES stocks during market crisis propose that investors who prefer ESG stocks are less sensitive to past performance. For example, Renneboog et al. (2011) provide evidence for the fact that investors in sustainable and responsible investment funds are less sensitive to past performance compared to others.

I use cross-sectional regression and multiple differences-in-differences regressions to test for the impact of the three ESG pillars on the returns during the COVID-19 market crash

focusing on the European markets. I find positive relationship between the Social score and returns during the market crash, defined as time period from February 24 to March 20. Even after controlling for various firm characteristics and other factors the results show statistically significant coefficient for the Social score variable. On the other hand, the Environmental score is not significant predictor of the returns during the market crash. And interestingly, the Governance score has negative impact on the returns, indicating that firm with good governance practices is expected to underperform other firms during the marker crash.

One of the interesting findings here is that only the Social score had positive impact on returns. Most studies have not found differences in the impact of the Environmental and Social scores (see Lins et al. (2016), Albuquerque et al. (2020), Ding et. al (2020), Demers et. al (2021)). Usually the Environmental and Social score are just aggregated to one score with the assumption that there is no difference in the impact for the two since they are highly correlated.

Statistically significant negative impact of the Governance score is also curious finding. Studies so far, mostly, has not found any meaningful impact for the Governance score during market crashes (see Lins et al. (2016), Koskinen (2019), Pastor and Vorsatz (2020)). Some studies have even recorded positive impact (see Lins et al. (2013) and Nguyen (2015)). These results hint that there might be more differences in the three ESG pillars in terms of their impact on returns during market crashes.

I also test whether the COVID-19 situation in country the firm is headquartered in affects the impact. The idea behind this is that the firm's located in countries with worse COVID-19 situation experience more severe exogenous shock than other firms, and thus I hypothesize that the impact of ESG policies should be more significant for these firms. Indeed, I find that impact of the Social score is positive and statistically significant for firms located in countries with worse COVID-19 situation measured by the case numbers during the market crash, whereas there is no significant impact found for other firms, i.e. firms located in countries better COVID-19 situation. Similarly, the Governance score has statistically significant negative impact only for firms in countries with relatively worse COVID-19 situation. These findings enforces the hypothesis that the impact of the Social and Governance scores on the

returns is due to the market crash, since the impact is only evident for firms experiencing harder shock.

Another hypothesis I test is how does firms with high Environmental, Social and Governance scores perform after the crisis. Nofsinger and Varma (2014), and Jurvanen (2020) provide empirical evidence in their studies for underperformance after the market downturn during the recovery period. They are essentially proposing that investor pays for the insurance provided by high ES stocks during the market shock by weaker subsequent performance. I do find some evidence for this effect but it is not as strong as the evidence found for the overperformance during the crisis. Firms with high social score and which are headquartered in countries that were hit harder by the pandemic are underperforming after the market crash. Since these firms were also outperforming during the crash, this indicates a risk-based explanation for this effect. In essence, this would mean that an investor could buy stock with high social score as a partial hedge, as the stock would be more resilient during market crashes, but the investor would have to pay for this hedge by the stock's subsequent underperformance.

Rest of this thesis is arranged in the following way. In section (2) I review the literature related to this study. Section (3) explains how I have constructed the hypotheses. Section (4) expounds on the data used. In section (5) I discuss the methods used. Section (6) encloses the results. In section (7) the robustness tests are performed. Section (8) discusses the outcomes. And finally, section (9) wraps up this work.

2 Literature review

In this section I will review the literature related to this thesis. First, I discuss the ESG and CSR measures and the problems that lie in measuring firm's social responsibility. Second, I review the literature that debates whether CSR/ESG has impact on firm performance and is the impact positive or negative. Third, I go to more specific topic discussing the impact of ESG on performance during market crisis. Finally, I review the literature discussing ESG during the COVID-19 pandemic.

2.1 ESG and CSR measures

The concept of corporate social responsibility (CSR) has a relatively long history. This history can be traced back to the 1930s and since then the concept of CSR has had varied history culminating in the current view of seeing CSR as a way to create shared value in society.[3] The alternative term ESG (environmental, social and governance) is replacing CSR as the leading term in use when discussing the larger role of firms in societies. Whereas CSR aims to make firms more accountable, ESG takes more wholly approach by focusing on measurability, transparency and integration of ESG into the business.[4] In this thesis I use CSR and ESG interchangeably unless otherwise stated.

There are multiple providers of ESG ratings. The most often used ESG raters in finance research studies seems to be KLD by MSCI and Refinitiv's ESG Scores provided by Thomson Reuters. Other, relatively well known, ESG raters are Sustainalytics, Vigeo-Eiris by Moody's and RobecoSAM by S&P Global. The problem is that the ratings from different providers can disagree meaningfully. Chatterji et al. (2016) document that the different ESG providers do not have same definitions of ESG, i.e. the raters differ in what they are trying to measure. In addition, Chatterji et al. (2016) find that even after adjusting for this difference in definition

[3] See Agudelo et. al (2019) for review of the history and evolution of the CSR concept

[4] See comparison of CSR and ESG in https://www.alva-group.com/blog/whats-the-difference-between-csr-and-esg/

there is substantial divergence in ratings between the raters. Dorfleitner et al. (2015), and Semenova and Hassel (2015) come to the same conclusion of divergent ESG ratings. More recent study by Berg et. al (2020) finds that in their dataset with six different providers of ESG ratings the average correlation between the ratings is 0.54. Berg et. al (2020) also tries to answer the question why ESG measurement diverge. According to them measurement divergence is the most important reason why the ratings diverge. This means that the raters measure the same firm in same category differently. Another important factor for the divergence effect, according to Berg et. al (2020), is how certain rating providers include some categories that others don't.

The well document fact that ESG ratings diverge between the raters demonstrate the subjectivity of these ratings and therefore sets some questions about the reliability of results in research that uses these ratings. I use Refinitiv's ESG scores in this thesis and the reader should be cognizant of the possibility that using ratings from different ESG rater could give different result.

2.2 Relationship between CSR and firm performance

Even though this thesis does not study the relationship between ESG and corporate financial performance directly, but rather the link between ESG and stock returns, it's important to get an idea how ESG activities affect the financial performance of the firm, since the financial performance of a firm and its stock returns are interlinked.

The relationship between CSR and firm performance is extensively researched topic but there is no clear consensus in which direction this relationship tilts, i.e. is there positive, negative or no relationship. Even if significant positive or negative relationship is found the causality of the effect is hard to determine. Still, currently the consensus view of the relationship between CSR and firm performance seems to be tilting towards positive effect. Friede et. al (2015) find positive relation between ESG and corporate financial performance (CFP) by aggregating results of 2000 empirical studies. Also, other review papers find evidence of positive relation between ESG or CSR and firm's financial performance. For

example, Malik (2015) and Clark et. al (2015) find support for the view of positive association between corporate financial performance and ESG activities.

If firm's ESG activities increase the firm's financial performance these can be expected to lead to increased firm value and thus higher stock returns. Clark et. al (2015) suggest three ways how ESG practices can lead to competitive advantage and therefore better business performance and ultimately higher stock returns. According to the authors the three ways are decreased risk, higher performance and better reputation. This suggest that firm can increase its market value by doing ESG activities. But this is different question from the one that asks if the ESG practices are priced in the stock price, i.e. do investors value the current state of firm's corporate social responsibility correctly.

Edmans (2011) suggests that intangible assets are not fully priced into the stock price. He finds evidence of higher longer term returns for firms with high levels of employee satisfaction. Similarly, Eccles et. al (2014) find that high sustainability corporations outperform low sustainability corporations during 18-year period. Servaes and Tamayo (2013) report that only high-CSR firms with high customer awareness (proxied by advertising expenditures) are positively related to firm market value. These papers support the view that investors are not fully valuing the benefits of ESG for firm's financial performance.

There is also research that documents evidence of lower cost of capital for firms with high CSR ratings. El Ghoul et. al (2011) find that especially high scores in the following CSR categories, employee relations, environmental policies and product strategies, have contributed most to the decrease of cost of equity capital. Also, Dhaliwal (2011) find that firm's that disclose voluntarily CSR activities with superior CSR performance enjoy reduction in the cost of capital. So, by doing CSR activities firms can lower the cost of equity capital and therefore increase firm value.

Besides these studies that demonstrate positive relationship between firm financial performance and ESG, there is also studies that report negative relationship. This negative relationship is often attributed to agency problems. The agency problems are born from the different incentives of the shareholders and the firm insiders. Masulis and Reza (2015) find that corporate giving is not purely about maximizing firm value since the CEOs have their

personal incentives. Similarly, Cheng et. al (2013) discover evidence of agency problems between managers and shareholders when it comes to investment in CSR: Marginal dollar spent on CSR does not maximize firm value. Also, Barnea and Rubin (2005) find CSR to be source of agency problems. On the other hand, Ferrell et. al (2016) find no clear evidence of agency problems by demonstrating that firms with good governance, and thus lower concern of agency problems, engage more in CSR activities. Therefore, firm's CSR practices would not be inconsistent with shareholder value maximization.

There is also a body of literature that suggest that there is no clear relationship between CSR and firm performance. Revelli and Viviani (2015) conduct a metastudy and provide evidence for the view that there is no real cost or benefit for including corporate social responsibility into investment decisions. According to them the findings depend on the the methodological choices made by the researchers. By using three different ESG rating concepts Halbritter and Dorfleitner (2015) do not find significant return differences between firms with high ESG ratings and firms with low ratings. Halbritter and Dorfleitner (2015), also, show that the effect of ESG scores on returns is influenced by the rating provider used (related to the literature about the divergence of ratings between the different raters, see above section (2.1)) and that the effect has decreased over the years. Renneboog et. al (2008) report that socially responsible investment (SRI) funds strongly underperform in large number of countries including US, UK, and many countries in continental Europe, but there is no statistically significant difference in risk-adjusted returns between SRI funds and conventional funds. Krüger (2015) finds that CSR is conditionally advantageous for firm value. His analysis shows that markets react positively to positive CSR news when it addresses problematic shareholder relations, i.e. when agency problems are a lesser issue. Without this condition investors react weakly negatively to positive CSR news.

One dimension that is relevant to this thesis is geographic location. Most of the studies discussed here, and on this topic overall, focuses on the US markets. But since this thesis is conducted using European data, it is important to compare the results across geographies. Friede et. al (2015) find in their meta-analysis that studies done with US data have larger share of positive results compared to studies done using European data. This suggests that the impact of ESG is larger in US markets than in European markets.

2.3 ESG and performance during market shocks

Empirical studies have found that firms with high CSR ratings have outperformed other firms during market shocks. Lins et al. (2017) find that firms with high CSR ratings perform better than low CSR firms during the 2008-2009 financial crisis by at least four percentage points while controlling for various risk factors and firm characteristics. Lins et al. (2017) also find higher profitability for high-CSR firms during the crisis. Nofsinger and Varma (2014) study the performance of socially responsible mutual funds and find that these funds outperform conventional funds during market downturns but underperform during other periods. These studies use US market data. Koskinen (2019) studies the European markets and finds positive and statistically significant relationship between high CSR score and returns during the 2008-2009 financial crisis.

Some theoretical explanations for the outperformance of high CSR firms during market crisis have been suggested. Lins et al. (2017) attribute this outperformance to social capital by showing that excess returns of high CSR firms are higher for firms headquartered in US regions with more trusting individuals. The claim is that during periods of crisis the trust in firms have deteriorated and therefore high social capital, created by CSR activities, will support the performance of high CSR firms during crises. Albuquerque et al. (2019) find that systematic risk is significantly lower for high CSR firms. They suggest that CSR activities increase product differentation by creating loyal customer base with less elastic demand. Therefore, the firm can attain higher profit margins, which lowers operating leverage and ultimately lowers the systematic risk. Crisis can be expectected to effect consumer demand and thus the customer loyalty can contribute to more robust stock performance during the crisis for high CSR firms.

There is also literature that suggest that investors who prefer high CSR stocks are not as sensitive to past returns as other investors. Renneboog et. al (2011) find that investors in SRI funds are less sensitive to past performance compared to investors in conventional mutual fund.[5] This suggests that investors with preference for CSR consider nonfinancial information in their decision making. Dooren and Galema (2018) provide evidence for higher dispostion effect in CSR investors. This means that socially responsible investor have tendency to hold

[5] See also Benson and Humphrey (2008)

the losing stocks and sell the winning stocks. These results indicate that during market crisis socially responsible investors are less likely to sell their socially responsible investments compared to other investors and thus high CSR stocks are more likely to perform better during market downturn.

Additionally, there exist body of literature that looks at the ability of ESG to decrease idiosyncratic risk, i.e. firm-specific risk that can be tied to legal or regulatory events. Godfrey et. al (2009) find that CSR can offer insurance-like protection during negative legal or regulatory event.[6] Hong et al. (2019) find evidence of lowered sanctions from prosecutors for ESG firms. Positive ESG practices and thus reduced idiosyncratic risk can be seen as more important during periods of crises since people pay more attention to bad coroprate behaviour during bad times and also it can be hypothesized that investors want to protect themselves from idiosyncratic risks during bear markets, and therefore leading to more robust performance for socially responsible firms during market crisis.

2.4 ESG and COVID-19

The recent COVID-19 crisis has provided new data to study the relationship between ESG and firm performance. There is already body of literature that has taken advantage of the 2020 COVID-19 crisis. For example, Ramelli and Wagner (2020) find that firms with low corporate debt and high cash holdings were able to weather the crash better.

Albuquerque et al. (2020) are, to my knowledge, the first to publish their findings on the effect of ESG on stock returns during the COVID-19 market shock. Albuquerque et al. (2020) find that stocks with higher environmental and social ratings had higher returns during the first quarter of 2020 in US markets. In addition, they observe lower volatility and higher operating margins for firms with higher environmental and social ratings.

Jurvanen (2020) finds evidence of high-CSR stocks outperforming during the COVID-19 market shock, but during the market recovery the high-CSR stocks underperform in US markets. This provides support for the Nofsinger and Varma's (2014) observation while

[6] See also Godfrey (2005)

studying the socially responsible funds, namely that overperformance during market downturns comes at the cost of underperformance during other periods. Jurvanen (2020) also provides evidence that the effect of better performance for socially responsible firms during the COVID-19 is not found or is statistically weaker for firms in STOXX Europe 600 index. In my study I use larger sample including all European firms with over 250M€ market cap and all necessary data and I also do more exhaustive research on the impact in the European market compared to Jurvanen (2020) who focuses more on the US markets in his study. I also study the individual impacts' of the three pillar scores, whereas Jurvanen (2020) focuses on the aggregate measure comprising of the Environmental and Social scores.

Using global dataset with over 6,000 firms Ding et. al (2020) find that firms with strong ESG policies before the COVID-19 pandemic performed better during the crisis in terms of stock returns. Demers et. al (2021) provide evidence to the contrary. The study claims to refute the previous findings of the significance of ESG as resiliency factor during the COVID-19 crisis by developing return model that includes control variables that other researchers have not considered, but which still is shown to be theoretically or empirically correlated with returns. On the other hand, Demers et. al (2021) find evidence for positive effect of intangible assets on the returns during the first quarter of 2020. Li et al. (2021), also, show that strong corporate culture, as an intangible asset, had positive effect on firm performance during the COVID-19 crisis. Findings by Folger-Laronde et al. (2020) show that higher sustainability of exchange-traded funds do not shield them from losses during a market crash. In the same vein, Döttling and Kim (2020) document that funds with highest sustainability rating reported by Morningstar experience sharper decline in net fund flows compared to other funds during the COVID-19 market crash. This challenge the evidence found before the COVID-19 crisis that ESG investors consider non-financial metrics more than past returns and thus the fund flows should be less volatile. (see e.g. Renneboog et. al (2011)). On the other hand, Pastor and Vorsatz (2020) find that high susainability equity mutual funds in US, especially funds focused on environmental sustainability, performed better than other funds during the crisis. Singh (2020) suggests, by studying portfolio performances during the COVID-19 pandemic, that the outperformance of high ESG stocks during crisis is due to investors becoming more attentive to corporate fundamentals and thus causing capital flows to ESG portfolio from EAFE and defensive portfolios during the crisis period.

There is also studies on COVID-19 and ESG that focus on certain geographical areas. For example, Palma-Ruiz et al. (2020) show that in Spain firms that made donations during the COVID-19 crisis performed better than other firms suggesting that during crisis firms that seem to be socially responsible can be expected to have better stock market performance compared to other firms. Evidence of high ESG outperformance is also found in stock markets in Far East. Broadstock et al. (2021) find that high-ESG portfolios outperform low-ESG portfolios in Chinese stock market during the COVID-19 shock, showing that ESG activities mitigate risk during financial crisis. On the other hand, Takashi and Yamada (2020) does not find evidence of higher performance for stocks with high ESG scores in Japanese stock markets.

As one can see consensus of whether firms with high ESG ratings outperformed other firms during the COVID-19 market crash has not yet been completely formed. Before going more deeply to my contribution to this literature, I'm taking little time to explain the spread of COVID-19 and its effect on financial markets and economy, and why this is good opportunity to study the impact of different variables on returns during market crashes. The COVID-19 crisis was global exogenous shock to markets that did not emerge from economic concerns, like the 2008 financial crisis, but rather from public health concerns. The fast and seemingly unexpected spread of the virus from China to the whole globe, and China's extreme measures to control the situation, and finally similar measures taken in Italy, scared investors which led to a market crash. The STOXX 600 index, measuring 600 largest firms from developed European countries, dropped by about 30% during a time period from February 24 to March 20. The rapid course of these events didn't left much time for firms to react to the situation, and therefore investors mostly reacted to the firm's condition before the crisis. These factors create opportune opportunity to study the effects of social responsibility on the returns during crisis.

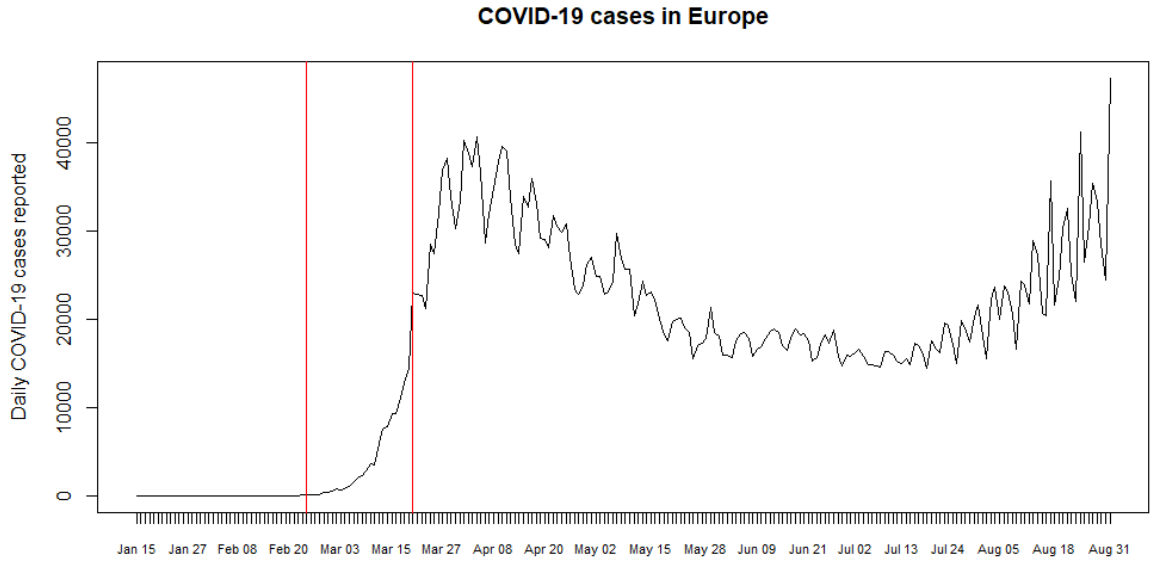

Figure 1 *Daily COVID-19 cases reported in Europe*
Aggregated data of reported daily COVID-19 cases of all European countries from January 15 to the end of October 2020. The horizontal red lines denote the defined start of the crash period February 24 and the end of the period defined at March 20. Data downloaded from European Centre for Disease Prevention and Control website: https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide

3 Hypotheses Construction

In this section I construct the hypotheses of this thesis based on theories and findings by previous studies. The main research question I am trying to answer is: does ESG performance of firm affect the stock returns during the COVID-19 market crash in Europe. The sole focus of this paper is on European markets. The studies on relationship between ESG and returns has mostly been focused on US markets, whereas the European markets have received less attention. Other research questions I try to answer are the following. I test if the effect of ESG is different during COVID-19 crash period compared to the period following the crash. One novel approach I take is to see if there is difference in the effect when taking into account how COVID-19 has affected European countries differently by separating firms based on the countries they are headquartered in. Finally, I ask whether the different components of the ESG score have different impact on the returns.

Next, I state each research questions and the related hypotheses one by one and provide justification for the hypotheses.

Q1: Does corporate social responsibility affect the stock returns during the COVID-19 market crash in Europe?

H1: Firms with high Environmental and Social scores outperform firms with low scores during the COVID-19 market crash.

H2: Firms with high Environmental and Social scores underperform firms with low ES rating during the COVID-19 market crash.

H0: There is no difference in stock returns between firms with high and low Environmental and Social scores during the COVID-19 market crash.

For the first hypothesis (*H1*) there is already body of empirical evidence. Lins et al. (2017) find that high ES (Environmental and Social) firms outperform other firms during the 2008-2009 financial crisis even after controlling for various firm characteristic and risk factors.

Similarly, Nofsinger and Varma (2014) find evidence of outperformance of socially responsible mutual funds during bear markets. Also, studies done already using data from the COVID-19 crisis support the first hypothesis. Albuquerque et al. (2020) find ES firms' stocks outperforming during the COVID-19 crash. Jurvanen (2020) also provides evidence of stock outperformance for the socially responsible firms. All of these studies are done using US stock market data, only Jurvanen (2020) tests the effect using European stock market data. Jurvanen (2020) does find only weak effect in the European markets.

There is also theoretical support for the first hypothesis. Lins et al. (2017) suggest that investors attach premium to firm's social capital during crisis and this social capital can be created by CSR activities. Albuquerque et al. (2019, 2020) theorize that socially responsible firms create loyal customer base and the creation of product differentation will help the firms perform financially better during crisis. Literature has also shown that socially responsible investors are less sensitive to past performance of socially responsible funds (see e.g. Renneboog et. al, 2011). The fact that socially responsible investors consider non-financial aspects in their investment decisions suggests that during crisis the fund outflow is not as large as for conventional investors. Also, idiosyncratic risk of high CSR firms has shown to be smaller compared to other firms (see Godfrey et. al (2009) and Hong et al. (2019)). During crisis legal and regulatory issues can possibly come to light more easily and therefore socially responsible firms are better postioned especially during crisis.

For the second hypothesis (*H2*) of underperformance, Jurvanen (2020) has two suggestions. First, Jurvanen (2020) suggests that high ESG firms may suffer from operational inefficiency by not taking part in optimal cost cutting practices that conventional firms are ready to take, and this leads to weaker financial performance and ultimately lower returns. The other suggestion by Jurvanen (2020) is that during crisis the increased attention toward ESG activities might reveal greenwashing done by high ESG firms and thus negatively affect the firm's stock performance during the crisis. This is also related to the agency problem literature that claims that manager's own interests might affect the investments in ESG practices (see e.g. Masulis and Reza (2015))

The null hypothesis (*H0*) suggests that all the relevant information of the ES score is included in the stock price and during the crisis possible effect of ESG on financial performance will

not be considered by investors. There is literature that claims that ESG has been irrelevant factor for investor's returns. Revelli and Viviani's (2015) metastudy documents that considering social responsibility in investing decisions do not affect the expected returns. Halbritter and Dorfleitner (2015) find that there are no significant return differences between high ESG and low ESG firms. Despite the studies by Albuquerque et al. (2020) and Jurvanen (2020) that find ESG to be relevant for returns in COVID-19 crisis, Demers et. al (2020) provide evidence for the fact that ESG does not have significant effect on returns during the COVID-19 market crash, and thus providing support for the null hypothesis.

Next, I provide brief explanations of the other research questions I try to answer:

Q2: Is the impact of ESG different during the crash compared to other time periods?

H1: After the crash firms with high Environmental and Social scores underperform the firms with low scores.

H2: After the crash firms with high Environmental and Social scores overperform the firms with low scores.

H0: After the crash there is no difference in the performance.

Lins et al. (2017) provide evidence in their paper that only during the crisis period the high ES firms outperform the low ES firms and during the periods surronding the crisis there was no significant difference in performance. Thus Lins et al. (2017) report evidence supporting the null hypothesis. On the other hand, Jurvanen (2020) and Nofsinger and Varma (2014) find that after the crisis high ESG firms underperform other firms. Basically, this means that ESG firms can provide insurance like benefits during the crash but the investor must pay for this by lower returns after the crash. The hypothesis that the high ES firms outperform other firms after crisis can be supported by the literature that has found overall evidence for outperformance for socially responsible firms (see literature review).

Q3: Is the relationship between the ESG scores and returns during the crash period different with firms headquartered in countries with more perilous Covid-19 situation?

H1: In countries with more perilous COVID-19 situation the difference in stock performance between high ES and low ES firms is more pronounced than in countries with more manageable COVID-19 situation.

H0: There is no difference in stock performance between high ES and low ES firms in countries with more perilous COVID-19 situation compared to countries with more manageable COVID-19 situation.

The idea behind the hypothesis is that the impact of ES on returns for firms located in countries with worse COVID-19 situation should be greater than for other firms is the following. First, countries where the virus started spreading rapidly during the start of the pandemic in February and March, e.g. Italy, experienced higher exogenous shock than other countries. This led to higher economic uncertainty. Therefore, I expect the impact of ES ratings on returns to be larger for these firms based on the assumption that these high-ES stocks operate as kind of insurance during market crisis as suggested by Lins et al. (2016).

What is the contribution of Environmental, Social and Governance score to the returns during the COVID-19 market crash?

Since I include all the three pillars of ESG in the regression independently I will also see how each of these contibute to the returns. Most past research has found that there is not much difference between the impact of the Environmental nad Social scores (see e.g. Lins et al. (2016), Albuquerque et al. (2020), Ding et. al (2020) and Demers et. al (2021)). Also, the correlation between the two measures is high (see e.g. Albuquerque et al. (2020)). So, the expectation, based on the past research and high correlation, is that there is not difference between the impact of the Environmental and Social scores, i.e. either both have impact or neither has.

The governance score is often left out of the analysis because it has not been kept as part of the firm's social responsibility, which usually taken as being the combination of firm's

environmental and social activities (see e.g. Lins et al. (2016)). I will incude the Governance score in the regression to avoid omitted variable bias and also see if the Governance score has any signicant impact on the returns. Past findings have often concluded that the governance score is not significant factor for returns during market crisis (see e.g. Lins et al. (2016), Koskinen (2019) and Pastor and Vorsatz (2020)). Based on these it can be expected that the governance score has no impact on returns during COVID-19 market crash.

4 Data and sample

In this section I explain the data I use. The data I need to test my hypotheses include ESG data, return data of the firms and firm characteristic data. The sample includes firms headquartered in Europe. ESG data is obtained from Thomas Reuters' Refinitiv ESG database. Firm characteristic and return data are obtained from Datastream. I first explain the data I use for the analysis and after that I demonstrate how the sample is constructed.

4.1 ESG data

The ESG data is collected from Thomas Reuters' Refinitiv ESG database. Refinitiv's ESG scores are based on public data and measure the ESG performance, commitment, and effectiveness across ten categories. These categories are divided into three ESG pillars: Environmental, Social and Governance. The Environmental pillar consists of the following categories: resource use, emissions and innovation. The Social pillar includes workforce, human rights, community and product responsibility categories. Management, shareholders and CSR strategy are the categories of the Governance pillar. Each of these ten categories consist of related themes. For example, the emission category includes the following four themes: emissions, waste, biodiversity, and environmental management system. The workforce category encloses four themes: diversity and inclusion, career development and training, working conditions, and health and safety. Within these themes are total of over 450 ESG metrics which are aggregated to 186 ESG measures that are used to calculate the ten category scores. The Environmental pillar and Social pillar scores are calculated using category weights that vary across industries. The weights are calculated based on the importance of each category and theme for each respective industry. The pillar scores are documented in percentages. The scores for Environmental and Social pillars are based on relative performance on sector level and for Governance pillar on country level. [7] Depending on the regression I will either use scores of all the three scores individually as independent

[7] For more info see ENVIRONMENTAL, SOCIAL AND GOVERNANCE (ESG) SCORES FROM REFINITIV (2020): https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf

variables in the regression or I will use measure constructed from Environmental and Social scores explained in the next section.

The Thomas Reuters' Refinitiv ESG database is also used by Albuquerque et al. (2020) and Jurvanen (2020). It is possible that using ESG data from different ESG data provider could result in different outcome. As I explained in the literature review (section 2.1) the ESG scores can diverge significantly between different ESG data providers.

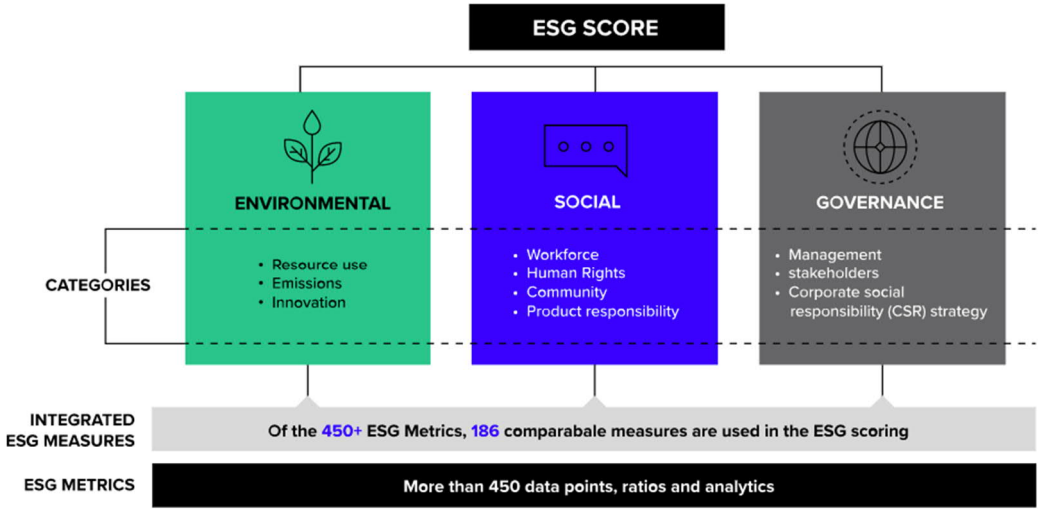

Figure 2 *ESG score*
Overview of what elements the ESG score is composed of. Figure taken from https://www.refinitiv.com/en/sustainable-finance/esg-scores

4.2 Stock return data and control variable data

The daily stock returns used in the regressions are obtained from Datastream. The return data is total return, that is, it also includes the dividends. Similarly, the firm characteristics I use as control variables are obtained from Datastream. The firm characteristics used are the following: *Cash Holdings* (cash and marketable securities divided by assets), *Profitability* (operating income divided by assets), *Short-Term Debt* (short-term debt divided by assets), *Long-Term Debt* (long-term debt divided by assets), *Size* (log of firm's market value of equity), *Book-to-Market* (book value divided by market value of equity), *Momentum* (12-month return prior to the crash period) and dummy for *Negative Book-to-Market* ratio (equals one if the book value is negative).

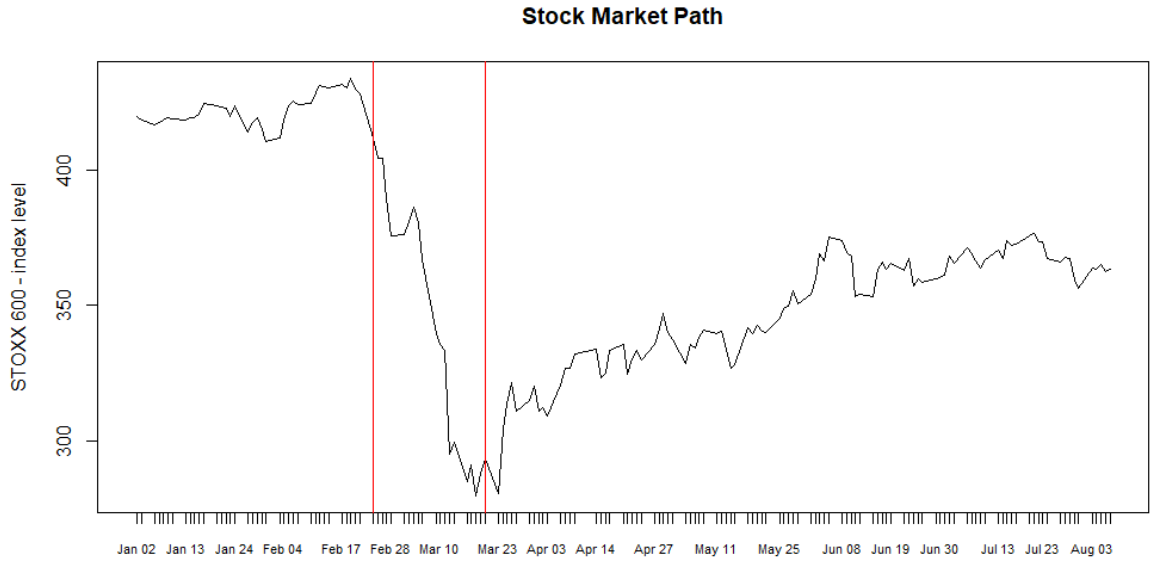

Figure 3 *STOXX 600 from the start of 2020 to the August of 2020*
The red lines denote the defined start of the crash period February 24 and the end of the period defined as March 20.

4.3 Sample construction

In this section, I first explain how I filtered the data due to missing data points. Second, I explain how I defined the COVID-19 crash period. Third, I explain how the abnormal returns are calculated.

I include all companies headquartered in Europe excluding the firms with following characteristics and missing data. I exclude firms with under 250€ million market capitalization at the end of 2019 as these micro-cap firms are likely to have low liquidity and high bid-ask spreads and these inefficiencies are likely to be even more pronounced during the market crash. This same exclusion is used by Lins et al. (2017). I exlude firms that do not have monthly return data from 2018 onwards. This quarentees that I have enough return data, i.e. two years, for market model regressions. Firms with missing Environmental, Social and Governance scores are exluded from the dataset. Lastly, I exlude firms that do not have the necessary firms characteristic data at the end of 2019. After these necessary exclusions I end up with 1104 firms. It is possible that these exlusions have created some biases in the sample. All of the sample data is winsorized at the 1st and 99th percentiles to avert problems

with outliers. This means that values that are smaller than the 1st percentile are set to the value of the 1st percentile, and all values larger than the 99th percentile are set to the value corresponding the 99th percentile.

I define the COVID-19 crash period from February 24 to March 20 of 2020. Ramelli and Wagner (2020) identify February 24 as the start of the 'Fever period'. On Sunday February 23, a lockdown of about 50,000 residents in towns located in northern Italy was announced. This was the first of multiple extraordinary events that followed in the subsequent weeks. The end of the defined crash period on March 20 precedes a Federal Reserve Board's announcement of major interventions in the corporate bond markets on March 23. Albuquerque et al. (2020) and Jurvanen (2020) define the period similarly with slight variation.

In addition to raw returns I also employ abnormal returns in the regressions. I use the Carhart (1997) Four Factor model with 5 years of monthly data. The model is based on the Fama and French three factor model, which includes the excess market retrun, size factor (SMB) and value factor (HML). The momentum factor (WML) is included as the fourth variable.

The factor data is obtained from Kenneth French's website. SMB stands for small minus big and is constructed by taking the average return of three small stock portfolios and subtracting it by the average return of three big stock portfolios. HML stands for high minus low and is calculated by taking the equal weight of two high book-to-market (value) portfolios and subtracting by the average of two low book-to-market (growth) portfolios. WML stands for winner minus loser, where the winners are the stocks with highest cumulative returns of the past year and losers with lowest returns. I calculate the cross-sectional regression to estimate the four parameters of the Four Factor model. I use these parameters to calculate the daily abnormal return.

Table 1 *Descriptive Statistics*

This table presents the summary statistics of the data used for the cross-sectional regression. Panel A provides the mean, standard deviation, median, 25th and 75th percentiles of the variables. Panel B shows the correlation matrix of the variables. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. 1104 firms are included in the sample. *ENV, SOC and GOV* show the environmental, social and governance pillar scores with the data from Refinitiv's ESG database as of year-end 2019. Data is winsorized at the 1st and 99th percentiles. *Crash-period raw return* is the return of the firm from February 24th, 2020 to March 20th, 2020. *Crash-period abn. return* is the abnormal return for the same time period; it is adjusted with Four Factor model estimated with 5 years of monthly data before the crash. Post-crash returns are calculated from March 20 to the end of August 2020. *Market cap* is the market capitalization of the firm in millions of euros at the end of 2019. The next six measures are also calculated from data available at the end of 2019. *Long-term debt* is equal to long-term debt divided by total assets. *Short-term debt* is measured by short-term debt and current portion of long-term debt divided by total assets. *Cash holdings* is equal to cash and short-term investments divided by total assets. *Profitability* is computed by dividing operating income by assets. *Book-to-market* is calculated by dividing the book value of the firm by the market value. *Negative B/M* is dummy variable that is equal to one when the book-to-market ratio is negative. *Momentum* is the 12-month return of the firm before start of the crash. *Idiosyncratic risk* is the residual variance of the market model.

PANEL A: Summary statistics

	Mean	Std Dev	25th perc.	Median	75th perc.
ENV	0.497	0.263	0.295	0.509	0.719
SOC	0.607	0.217	0.451	0.624	0.786
GOV	0.524	0.225	0.340	0.529	0.704
Crash-period raw return	-0.305	0.152	-0.406	-0.304	-0.204
Crash-period abn. return	-0.045	0.185	-0.162	-0.044	0.069
Post-crash raw return	0.464	0.421	0.170	0.386	0.676
Post-crash abn. return	0.094	0.448	-0.162	0.035	0.307
Market cap (M€)	9363	22498	874	2535	7529
Long-term debt	0.221	0.157	0.096	0.205	0.315
Short-term debt	0.052	0.055	0.013	0.036	0.073
Cash holdings	0.132	0.134	0.050	0.094	0.159
Profitability	0.075	0.087	0.040	0.068	0.106
Book-to-market	0.488	0.405	0.209	0.381	0.662
Negative B/M	0.021	0.143	0.000	0.000	0.000
Momentum	-0.238	0.266	-0.418	-0.256	-0.077
Idiosyncratic risk	0.008	0.007	0.003	0.005	0.009

PANEL B: CORRELATION MATRIX

	ENV	*SOC*	*GOV*	*Crash-period raw return*	*Crash-period abn. return*	*Post-crash raw return*	*Post-crash abn. return*	*Market cap (M€)*	*Long-term debt*	*Short-term debt*	*Cash holdings*	*Profitability*	*Book-to-market*	*Negative B/M*	*Momentum*
SOC	0.744														
GOV	0.431	0.459													
Crash-period raw return	0.015	0.034	-0.065												
Crash-period abn. return	0.041	0.098	-0.076	0.581											
Post-crash raw return	-0.164	-0.155	-0.068	-0.245	-0.134										
Post-crash abn. return	-0.153	-0.167	-0.038	-0.107	-0.409	0.89									
Market cap (M€)	0.512	0.5	0.398	0.138	-0.008	-0.173	-0.099								
Long-term debt	0.106	0.111	0.121	-0.104	-0.084	-0.061	-0.041	0.039							
Short-term debt	0.096	0.066	-0.02	0.036	0.062	-0.037	-0.042	-0.045	0.105						
Cash holdings	-0.256	-0.159	-0.124	0.069	0.147	0.119	0.041	-0.111	-0.279	-0.201					
Profitability	0.045	0.013	0.059	0.029	-0.106	0.003	0.061	0.193	-0.071	-0.057	-0.152				
Book-to-market	0.139	0.034	0.02	-0.139	0.042	-0.065	-0.119	-0.128	-0.048	0.04	-0.205	-0.265			
Negative B/M	0.014	0.021	0.004	-0.067	-0.006	-0.061	-0.076	-0.064	0.288	0.02	0.002	0.017	-0.222		
Momentum	-0.051	-0.016	-0.08	0.681	0.327	-0.079	0.026	0.159	-0.086	-0.014	0.115	0.068	-0.237	-0.031	
Idiosyncratic risk	-0.315	-0.273	-0.201	-0.102	0.074	0.227	0.105	-0.394	-0.04	-0.041	0.389	-0.241	-0.041	0.175	-0.018

4.4 Sample description

In this section I look at the sample data and its attributes. The panel A of table 1 provides the mean, standard deviation, median, 25th and 75th percentiles of the main variables used in the regression models. The Environmental (*ENV),* Social (*SOC*) and Governance (*GOV*) scores varies from 0 to 1. The mean of the Environmental score, is 0.497 and the median is 0.509. The mean of the Social score, is 0.607 and the median is 0.624. And respectively, the mean of the Governance score, is 0.524 and the median is 0.529. These means and medians suggest that, according to Refinitiv, European firms are especially socially responsible compared to firms outside Europe, since the Social score is substantially over 0.5. The *Crash-period raw return* is strongly negative, as expected, with mean of -0.305 and 25th percentile of -0.406. This is almost as large crash as Albuquerque et al. (2020) and by Jurvanen (2020) reported of their US samples. The mean of the *Crash-period abnormal return* is -0.045 with median of -0.044. On the other hand, the *Post-crash raw return* is highly positive with mean of 0.464. The panel A of table 1 also provides the statistics for the main control variables used in the regressions.

Panel B of table 1 provides the correlation matrix of the variables. From this matrix one can observe if some pairs of variables have high correlation. This can cause problems of multicollinearity, meaning that statistical significance of the variable could be undermined. There seems to be two appropriately higher correlations. First, the Environmental score and Social score has correlation of 0.744. The correlation between Environmental and Governance score is lower, 0.431. And respectively, the correlation between Social and Governance score is 0.459. Looking at the other variables, the correlation between market cap and the Environmental and Social scores is slightly over 0.5 for both, suggesting that larger firms are more socially and environmentally responsible than smaller firms. The other higher correlation is between the momentum variable and the raw return during the crash period. This correlation is high, 0.681, and shows that stocks that performed well during 12-months prior to the COVID-19 crash also endured the crash better, whereas stocks that underperformed during the prior 12-months also underperformed during the crisis.

5 Methods

The goal of this thesis is to study the relationship between the ESG scores and stock returns during the COVID-19 market crash and the period after it. This section provides information about the methods used. I use cross-sectional regressions as well as a differences-in-differences regressions. I apply the differences-in-differences regression to detect more clearly the effect of COVID-19 market crash on the relation between ESG and the stock returns. I use the three scores of Environmental, Social and Governance individually in the regressions. The methods I employ in my thesis are inspired by Albuquerque et al. (2020). Also, Lins et al. (2017) and by Jurvanen (2020) apply similar methods in their studies.

5.1 Cross-sectional model

5.1.1 Cross-sectional regression

For the cross-sectional regression I define the time period of the crash from 24.02.2020 to 20.3.2020. I use two different dependent variables *Raw Crash Return* and *Abnormal Crash Return. Raw Crash Return* is equal to the return during the crash period. The *Abnormal Crash Return* is equal to raw return minus the expected return. The expected return is calculated using the Four Factor model with 60 months of return data before the crash period. (To see more minute explanation of the construction of the abnormal returns see section 4.4.) I include the Environmental, Social and Governance scores individually as independent variables in the regression. These are the important variables that are in the spotlight.

I include number of control variables in the regression. I control for firm characteristics that could affect the firm performance during the crash. These firm characteristics control variables are similar to the ones Lins et al. (2017) use. The firm characteristic control variables that are included are *Cash Holdings* (cash and marketable securities divided by assets), *Profitability* (operating income divided by assets), *Short-Term Debt* (short-term debt divided by assets) and *Long-Term Debt* (long-term debt divided by assets). These firm characteristics can meaningfully impact the firm's performance during a market crash. During crisis firms with large cash balances, moderate leverage and profitable business are

more likely to be able to continue run their business during turbulent times (see e.g. Harford et. al (2012)). Ramelli and Wagner (2020) show that during the recent COVID-19 shock cash holdings and corporate debt were important factors for firm value.

In addition, I include firm characteristics that have been shown to affect the stock returns (see Daniel and Titman (1997)). This includes *Size* (log of firm's market value of equity), *Book-to-Market* (book value divided by market value of equity), *Momentum* (12-month return prior to the crash period) and dummy for *Negative Book-to-Market* ratio, i.e. the dummy is equal to one when the book-to-market value is negative. I also include *Idiosyncratic Risk,* which is computed as the residual variance of the market model. The idiosyncratic risk controls for the possibility that price volatility affects the returns.

In addition to these control variables, I also include the following fixed effects in the regressions. First, industry fixed effects are included. The industry fixed effects are based on The Refinitiv Business Classification (TRBC) industry classification system. I use the highest-level codes where firms are broken up to 13 economic sectors. Industry fixed effects are included because COVID-19 has affected industries differently, as Ramelli and Wagner (2020) has observed. Unsurprisingly, firms that operate in industries where social distancing and lock-down measures are relevant obstructions for conducting business in normal fashion have experienced larger disruptions in their operations (see also Pagano et al. (2020)). In addition to industry fixed effects, I include country fixed effects based on the country of exchange. It is clear that COVID-19 has had different impact in different European countries. For example, Italy was already in nationwide quarantine during March, whereas in northern European countries the restriction measures were more lenient.

We end up with the following cross-sectional regression:

$$Crash\ Return_i = \beta_0 + \beta_1 ENV_i + \beta_2 SOC_i + \beta_3 GOV_i + \beta_4 Firm\ Controls_i + Industry\ FE_i + Country\ FE_i + \varepsilon_i.$$

To see the results of this model see section (6.1)

5.1.2 Cross-sectional regression with dummies for quartiles

I also test another cross-sectional regression where I divide the firms to quartiles based on their ES score. Similar test is done by Lins et al. (2017). So, instead of using the linear ESG scores as independent variable, this model has three dummy variables for each Environmental, Social and Governance quartiles. The first quartile is captured by the intercept in each three cases. A firm is included in one of the four quartiles based on how the firm ranks relation to other firms in terms of its Envirnomental, Social or Governance score. The corresponding dummy variable equals one and the other dummies are set to zero for the particular firm. With this model it is possible to see if the impact on stock returns is stronger for firms with very high score. The control variables are the same as in the baseline cross-sectional regression above. This cross-sectional model looks as follows:

$$Crash\ Return_i = \beta_0 + \beta_1 ENV2_i + \beta_2 ENV3_i + \beta_3 ENV4_i + \beta_4 SOC2_i + \beta_5 SOC3_i + \beta_6 SOC4_i + \beta_7 GOV_i 2 + \beta_8 GOV3_i + \beta_9 GOV4_i + \beta_4 Firm\ Controls_i + Industry\ FE_i + Country\ FE_i + \varepsilon_i.$$

Where, for example, $ENV2_i$, $ENV3_i$ and $ENV4_i$ are the dummy variables for the Environmental score quartiles, $ENV4_i$ being the dummy for firms in the quartile with highest Environmental scores. Same case for the Social and Governance dummy variables. ENV1, SOC1 and GOV1 are captured by the intercept term.

To see the results of this model see section (6.2).

5.2 Differences-in-differences model

5.2.1 Differences-in-differences regression

I apply the differences-in-differences regressions to test whether there are differences in the relation of ESG and stock returns during COVID-19 crash and time periods surrounding it. The sample period is from August 2019 to August 2020, it starts before the crash and ends after the crash. Daily data is used to construct the panel. The differences-in-differences model is as follows:

$$Return_{i,t} = \beta_0 + (\beta_1 ENV_{i,2019} + \beta_2 SOC_{i,2019} + \beta_3 GOV_{i,2019}) * Crash_t + (\beta_4 ENV_{i,2019} + \beta_5 SOC_{i,2019} + \beta_6 GOV_{i,2019}) * Post\text{-}Crash_t + Time\ FE_t + Firm\ FE_i + \varepsilon_{i,t}.$$

The $Return_{i,t}$ is the daily raw return or the daily abnormal return (constructed similarly as in the cross-sectional regression). The variables $ENV_{i,2019}$, $SOC_{i,2019}$ and $GOV_{i,2019}$ are dummy variables for the Environmental, Social and Governance activity of the firm, where the variable is equal to one if the Environmental, Social or Governance score is in the top quartile. For example, if firm *i* has Social score that is in the top quartile, the $SOC_{i,2019}$ variable for the firm *i* is set to one. The ESG data used is acquired from the Refinitiv's database at the end of 2019, as in the cross-sectional regressions. $Crash_t$ is dummy variable that is set to one during the days from February 24 to March 20. $Post\text{-}Crash_t$ variable equals one for the days after the crisis from March 21 until the end of August. The differences-in-differences model includes firm fixed effects to control for omitted variables. Time fixed effects are also included on daily level to control for any time-series patterns. The coefficients $\beta_1, \beta_2\ and\ \beta_3$ capture the impact of the Environmental, Social and Governance scores on the returns during the crisis after controlling for the mentioned factors. The coefficents $\beta_4, \beta_5\ and\ \beta_6$ reflects the impact of the three different ESG scores on the returns after the crisis.

5.2.2 Regional Covid-19 situation and differences-in-differences regression

Next, I'm going to test whether there are differences in the relation between the three ESG scores and stock returns based on the seriousness of the Covid-19 pandemic in the country where the company is headquartered. The aim of this regression is to use the exogenous COVID-19 crisis to test whether the impact of the ESG scores on the returns is different depending on the uncertainty of the economic and social situation. Countries with more perilous COVID-19 situation are more likely to take harder measures to prevent the spread of the virus, and thus affecting the business performance of the firm operating there. Therefore, it can be hypothesized that investors took the corona situation into account when deciding which firms to buy by following the development of the reported COVID-19 case numbers in the country the firm is based on. Like above I use the Environmental, Social and Governance top quarter dummy variables.

In essence, I divide countries to high-covid and low-covid countries based on their reported cumulative COVID-19 case numbers during the months of February and March.[8] The cumulative COVID-19 case numbers are made proportional by dividing them by the population of the country. The splitting is based on the median calculated from the cumulative COVID-19 case numbers of the European countries, e.g. if a firm is headquartered in Finland and the cumulative Covid-19 case number of Finland is smaller than the median then the firm is included in the low-covid group. The model is as follows:

$$Return_{i,t} = \beta_0 + (\beta_1 ENV_{i,2019} + \beta_2 SOC_{i,2019} + \beta_3 GOV_{i,2019}) * Crash_t * Low\text{-}Covid_i + (\beta_4 ENV_{i,2019} + \beta_5 SOC_{i,2019} + \beta_6 GOV_{i,2019}) * Crash_t * High\text{-}Covid_i + (\beta_7 ENV_{i,2019} + \beta_8 SOC_{i,2019} + \beta_9 GOV_{i,2019}) * PostCrash_t * Low\text{-}Covid_i + (\beta_{10} ENV_{i,2019} + \beta_{11} SOC_{i,2019} + \beta_{12} GOV_{i,2019}) * PostCrash_t * High\text{-}Covid_i + Time\ FE_t + Firm\ FE_i + \varepsilon_{i,t}.$$

Where $High\text{-}Covid_i$ is set to one if the firm is headquartered in a country with relatively worse Covid-19 situation during the months of February and March. The dummy variable $Low\text{-}Covid_i$ is equal to one for firms headquartered in countries with lower Covid-19 case numbers during February and March. ENV, SOC and GOV are dummy variables set to one if the firm's score is in the top quartile. Time and firm fixed effects are included.

Here are some examples what the coefficients tell us. The coefficient β_1 captures the impact of the Environment score on the returns during the crash for firms headquartered in country with relatively less COVID-19 cases. β_2 reflects the impact of the Social score variable on the returns during the crash for firms headquartered in low-covid countries. The coefficient β_6, for instance, catches the impact of the Governance score on the returns during the crash for high-COVID countries. And one more example, the coefficient β_{10} captures the impact of the Environmental score on the returns after the crisis for firms headquartered in high-covid countries, and so on.

[8] The historical data for the COVID-19 cases reported in Europe are downloaded from the European Centre for Disease Prevention and Control website: https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide

6 Results

In this part I discuss the results I have estimated using the methods explained in part 5. First, I report the result of the cross-sectional regressions, and after that I describe the results I have obtained using differences-in-differences regressions. I leave the more meticulous discussion of these results for separate section.

6.1 Results of the cross-sectional regression

Table 2 provides the results of the cross-sectional regression. See the methodology used in section 5.1.1. The columns (1) and (2) presents the results where they are controlled only for country and industry fixed effects. In columns (3) and (4) the regressions include the control variables mentioned in the section 5.1. In columns (1) and (3) the dependent variable is the raw return during the crash period (from February 20 to March 24), whereas in columns (2) and (4) the dependent variable is the abnormal return during the crash period.

For the three main variables the only statistically significant positive coefficient is that of the Social score in regression where the independent variable is the abnormal return (columns 2 and 4). The positive coefficient suggests that firm's positive social activities before the crash could help its stock price during market crash. On the other hand, the coefficient of the Governance score is negative and statistically significant for the regressions with no control variables (columns 1 and 2), but this significance disappears when the control variables are included. The coefficient for the Environmental score is not statistically significant in any of the four regressions. Based on this there does seem to be some evidence of higher returns for firms with higher Social scores during the COVID-19 market crash.

Looking at the control variables it seems that at least the past 12-month returns (momentum) have had significant positive impact for the returns during the market crash. Also, cash holdings and book-to-market ratio had positive impact in the regression with abnormal return as independent variable. Interestingly, the leverage of the firm did not seem to affect the stock performance during the crash. See appendix for the results of a cross-sectional regression with the aggregated ES measure as the independent variable.

Table 2 *Cross sectional regression: Relation between ESG and stock returns*

This table presents the results of a cross-sectional regression estimating the relation between the ESG scores and stock returns during the COVID-19 market crash. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The model is as follows:

$$Crash\ Return_i = \beta_0 + \beta_1 ENV_i + \beta_2 SOC_i + \beta_3 GOV_i + \beta_4 Firm\ Controls_i + Industry\ FE_i + Country\ FE_i + \varepsilon_i.$$

The dependent variable is either raw return or abnormal return (based on the Four Factor model) from February 24 to March 20. The independent variables include the Environmental, Social and Governance scores and control variables. Columns 1 and 2 do include only the industry fixed effects and country fixed effects as control variables. The columns 3 and 4 add various firm characteristics as control variables (data from the end of 2019). These include Cash Holdings (cash and marketable securities divided by assets), Profitability (operating income divided by assets), Short-Term Debt (short-term debt divided by assets) and Long-Term Debt (long-term debt divided by assets), Size (log of firm's market value of equity), Book-to-Market (book value divided by market value of equity), Momentum (12-month return prior to the crash period), dummy for Negative Book-to-Market ratio and Idiosyncratic Risk (residual variance of the market model). The Heteroskedasticity-consistent standard errors are presented in parentheses. All data winsorized at the 1st and 99th percentiles. * p < 0.1, ** p < 0.05, *** p < 0.01

	Raw return	Abnormal return	Raw return	Abnormal return
	(1)	(2)	(3)	(4)
ENV	0.001	-0.035	0.004	-0.012
	(0.028)	(0.033)	(0.021)	(0.032)
SOC	0.050	0.105***	0.018	0.114***
	(0.032)	(0.039)	(0.027)	(0.039)
GOV	-0.046**	-0.069**	-0.019	-0.031
	(0.023)	(0.027)	(0.017)	(0.027)
Size			-0.001	-0.004
			(0.003)	(0.005)
Long-Term Debt			-0.036	-0.037
			(0.023)	(0.034)
Short-Term Debt			0.021	0.062
			(0.060)	(0.090)
Cash holdings			0.027	0.157***
			(0.038)	(0.057)
Profitability			-0.004	0.020
			(0.049)	(0.074)
Book-to-Market			0.003	0.076***
			(0.011)	(0.016)
Negative B/M			-0.026	0.037
			(0.032)	(0.053)
Momentum			0.357***	0.244***
			(0.017)	(0.023)
Idiosyncratic Risk			-2.367***	2.814**
			(0.713)	(1.217)
N	1104	1104	1104	1104
Adj. R^2	0.216	0.198	0.557	0.326
Country FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes

6.2 Results of the cross-sectional regression with dummies for quartiles

Table 3 represent the results of a regression where the ESG scores are divided to quartiles. This is done for all three components of ESG individually, namely Environmental, Social and Governance scores each get their own quartiles. The model has three dummy variables for the quartiles 2, 3 and 4 for all three ESG pillars. In each case the first quartile is captured by the intercept. *ENV4, SOC4* and *GOV4* include the firms with the highest scores. Columns (3) and (4) include same firm control variables as in the cross-sectional regression above. In columns (1) and (3) the dependent variable is the raw return and in columns (2) and (4) the dependent variable is the abnormal return during the crash period. For more detailed explanation see section 5.1.2.

The second and third quartiles of the Environmental score captures statistically significant coefficients in most cases, but the second quartile has positive coefficients and the third quartile negative coefficients for all cases. The top quartile is negative in all four regressions but not statistically significant in any. For the social score the second quartile is negative for all and statistically significant for all but the third regression. For the top two quartiles the coefficients are all positive and mostly statistically significant. For the Governance quartiles only the top quartile has statistically significant coefficients, these are found for the regressions without the firm control variables and both of them are negative.

These findings using the quartiles support the results found in the cross-sectional regression with linear ESG measures. Nothing definitive can be said about the impact of Environmental score, Social score seems to have positive effect on returns, and the Governance score seems to be negatively related to stock returns during market crisis.

Table 3 *Cross sectional regression: Relation between ESG and stock returns with dummies for quartiles*

This table presents the results of a cross-sectional regression estimating the relation between ESG pillar scores and stock returns during the COVID-19 market crash. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The model is as follows:

$$Crash\ Return_i = \beta_0 + \beta_1 ENV2_i + \beta_2 ENV3_i + \beta_3 ENV4_i + \beta_4 SOC2_i + \beta_5 SOC3_i + \beta_6 SOC4_i + \beta_7 GOV_i 2 + \beta_8 GOV3_i + \beta_9 GOV4_i + \beta_{10} Firm\ Controls_i + Industry\ FE_i + Country\ FE_i + \varepsilon_i.$$

The dependent variable is either raw return or abnormal return (based on the Four Factor model) from February 24 to March 20. $ENV2_i$, $ENV3_i$ and $ENV4_i$ are the dummy variables for the Environmental quartiles. Similarly for the Social and Governance quartiles. The fourth quartile includes the firms with the highest scores. Otherwise, this model is similar to table 2's model, i.e. it has the same control variables. Columns 1 and 2 do include only the industry fixed effects and country fixed effects as control variables. The columns 3 and 4 add various firm characteristics as control variables. The Heteroskedasticity-consistent standard errors are presented in parentheses. All data winsorized at the 1st and 99th percentiles. * p < 0.1, ** p < 0.05, *** p < 0.01

	Raw return	Abnormal return	Raw return	Abnormal return
	(1)	(2)	(3)	(4)
ENV2	0.019	0.046***	0.012	0.030*
	(0.014)	(0.017)	(0.011)	(0.016)
ENV3	-0.024*	-0.030*	-0.015*	-0.029**
	(0.013)	(0.015)	(0.009)	(0.014)
ENV4	-0.014	-0.007	-0.008	-0.006
	(0.012)	(0.015)	(0.010)	(0.014)
SOC2	-0.024*	-0.039***	-0.013	-0.042***
	(0.013)	(0.015)	(0.009)	(0.014)
SOC3	0.029**	0.015	0.020*	0.003
	(0.013)	(0.018)	(0.010)	(0.017)
SOC4	0.026**	0.018*	0.005	0.032**
	(0.013)	(0.016)	(0.010)	(0.013)
GOV2	0.013	-0.006	-0.012	-0.023
	(0.012)	(0.016)	(0.010)	(0.015)
GOV3	-0.019	0.007	0.007	0.013
	(0.012)	(0.014)	(0.009)	(0.013)
GOV4	-0.037***	-0.032**	-0.013	-0.017
	(0.012)	(0.014)	(0.009)	(0.013)
Firm controls	No	No	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
N	1104	1104	1104	1104
Adj. R^2	0.200	0.172	0.539	0.296

6.3 Results of the differences-in-differences regression

Table 4 contains the results of the differences-in-differences regression. The goal of this regression is to test what is the impact of the Environmental, Social and Governance scores during the crash and after it using event dummies $Crash$ and *Post-Crash* that determine each period on daily level. I use dummy variables to determine the Environmental, Social and Governance activity of the firm, where the variable is set to one if the score is in the top quartile. Column (1) shows results with the raw return as the dependent variable and column (2) has the result of a regression with the abnormal return as the dependent variable. Both include firm and time fixed effects. Standard errors are clustered by firm. To see more detailed explanation of the models see section 5.2.1.

Table 5 shows that the during crisis the coefficient for the Social score variable is positive and statistically significant at 1% level for both regressions. On the other hand, variable for the Governance score has negative coefficient which is also statistically significant at 1% level during the crisis period in both regressions. The coefficient for the Environmental score is not statistically significant during the crisis period. None of the coefficients for the ESG variables that measure the impact after the crisis are statistically significant. It still can be noted that the Social score variable has negative term which is almost statistically significant at 10% level.

These results support the evidence found for the cross-sectional regressions. Namely, that the social activities of firm before the COVID-19 crisis had positive effect on returns during the COVID-19 market crash. In effect, these results show that social activities could provide resiliency for firm's stock price during market crashes and the evidence that investor have to pay for this by underperforming after the crisis is weak. As in the cross-sectional regression, the differences-in-differences regression provide evidence for the fact that high Governance score has negative impact on the returns during the crisis.

Table 4 *Differences-in-differences regression: Relation between ESG and stock returns*

This table presents the results of a differences-in-differences regression estimating the impact of the ESG pillar scores on the stock returns during and after the COVID-19 market crash. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The sample period is from August 2019 to August 2020, starting before the crash and ending after the crash. Daily data is used. The model is:

$$Return_{i,t} = \beta_0 + (\beta_1 ENV_{i,2019} + \beta_2 SOC_{i,2019} + \beta_3 GOV_{i,2019}) * Crash_t + (\beta_4 ENV_{i,2019} + \beta_5 SOC_{i,2019} + \beta_6 GOV_{i,2019}) * Post\text{-}Crash_t + Time\ FE_t + Firm\ FE_i + \varepsilon_{i,t}$$

The dependent variable is either raw return or abnormal return (based on the Four Factor model) measured on monthly basis. $ENV_{i,2019}$, $SOC_{i,2019}$ and $GOV_{i,2019}$ are the Environmental, Social and Governance dummy variables equal to one if the firm's score is in the top quartile. The ESG data is from Refinitv's database, measured at the end of 2019. $Crash_t$ is dummy equal to one during the days between February 24 and March 20. $Post\text{-}Crash_t$ is dummy set to one during the time period after the crisis from March 21 to the end of August. Control variables include the firm fixed effects, and time fixed effects on daily basis. All data is winsorized at the 1st and 99th percentiles. The Heteroskedasticity-consistent standard errors are presented in parentheses. The standard errors are clustered at firm level.
* p < 0.1, ** p < 0.05, *** p < 0.01

	Raw return *(1)*	Abnormal return (2)
*ENV * Crash*	0.00623	0.01003
	(0.08318)	(0.08346)
*SOC * Crash*	0.33826***	0.33650***
	(0.08464)	(0.08490)
*GOV * Crash*	-0.26628***	-0.26880***
	(0.08073)	(0.08098)
*ENV * Post-Crash*	-0.01113	-0.00258
	(0.02596)	(0.02734)
SOC Post-Crash*	-0.04180	-0.04576
	(0.02707)	(0.02838)
*GOV * Post-Crash*	0.01559	0.00993
	(0.02538)	(0.02642)
Time FE (daily)	Yes	Yes
Firm FE	Yes	Yes
SE clustered by	Firm	Firm
N	287040	287040
Adj. R2	0.251	0.435

6.4 Results of the differences-in-differences regression with regional COVID-19 situation

Table 5 provides the results of the regressions that divides firms based on the country they operate from. For firms that are headquartered in countries with relatively more serious COVID-19 situation, based on the reported cases, the dummy variable $High\text{-}Covid$ is set to one. Other firms have the dummy variable $Low\text{-}Covid$ equal to one. I use the same Environmental, Social and Governance dummy variables in the regression. To see further explanation of the model see 5.2.2.

As the table 5 shows only coefficient for the firms located in high-COVID countries are statistically significant. This suggests that the size of the shock is important factor in the impact of the ESG scores. Again, the Social score is the positive and statistically significant factor in the regressions during the crash, but now only for the firms located in countries with worse COVID-19 situation. Now, the Social score's coefficient after the crash is negative and statistically significant for the high-COVID firms, supporting the risk-based explanation for the effect, i.e. an investor can hedge his portfolio returns during market crash by owning stocks with high social score but he have to pay for this by underperformance after the crisis. And again, high Governance score has negative impact on the returns during the crash, but only for high-COVID firms.

Table 5 *Regional Covid-19 situation and ESG in differences-in-differences regression*

This table presents the results of a differences-in-differences regression estimating the relation between ESG and stock returns before, during and after the COVID-19 market crash taking into account the COVID-19 situation during the crash. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The sample period is from August 2019 to August 2020, starting before the crash and ending after the crash. Daily data is used. See the model in the methods part. The dependent variable is either raw return or abnormal return (based on market model) measured on monthly basis. ENV, SOC and GOV are the Environmental, Social and Governance dummy variables that are equal to one if the firm's score is in the top quartile. The ESG data is from Refinitv's database, measured at the end of 2019. $Crash_t$ is dummy equal to one during the days between February 24 and March 20. $Post\text{-}Crash_t$ is dummy set to one during the time period after the crisis from March 21 to the end of August.$High\text{-}Covid_i$ is dummy variable that equals one if the firm is from country with severe

COVID-19 situation during the crash. $Low\text{-}Covid_i$, on the other hand, is equal to one when the firm is based in a country with relatively good COVID-19 situation. Control variables include the firm fixed effects, and time fixed effects on monthly basis. All data is winsorized at the 1st and 99th percentiles. The Heteroskedasticity-consistent standard errors are presented in parentheses. The standard errors are clustered at firm level.
* p < 0.1, ** p < 0.05, *** p < 0.01

	Raw return *(1)*	Abnormal return (2)
*ENV * Crash * Low-* $Covid$	-0.12894	0.00304
	(0.13590)	(0.11675)
*SOC * Crash * Low-* $Covid$	0.20359	0.12799
	(0.15341)	(0.14421)
*GOV * Crash * Low-* $Covid$	0.04720	0.02809
	(0.15816)	(0.15610)
*ENV * Crash * High-* $Covid$	-0.07647	0.11499
	(0.11539)	(0.11406)
*SOC * Crash * High-* $Covid$	0.40103***	0.27662**
	(0.11694)	(0.11531)
*GOV * Crash * High-* $Covid$	-0.28330***	-0.36543***
	(0.10026)	(0.09839)
*ENV * Post-Crash * Low-* $Covid$	-0.00204	-0.03162
	(0.09857)	(0.09670)
*SOC * Post-Crash * Low-* $Covid$	0.04211	0.05906
	(0.11579)	(0.11463)
*GOV * Post-Crash * Low-* $Covid$	0.03419	0.03848
	(0.13266)	(0.13241)
*ENV * Post-Crash * High-* $Covid$	0.08956	0.04665
	(0.07288)	(0.07260)
*SOC * Post-Crash * High-* $Covid$	-0.20259***	-0.17470**
	(0.07115)	(0.07079)
*GOV * Post-Crash * HighCovid*	0.01474	0.03315
	(0.07818)	(0.07791)
Time FE (daily)	Yes	Yes
Firm FE	Yes	Yes
SE clustered by	Firm	Firm
N	287040	287040
Adj. R2	0.255	0.437

7 Robustness tests

I conduct number of further tests to determine the robustness of the findings. First, I test regressions where I include only one of the ESG pillars in the regression. In the second setup I use different ESG variables. Thirdly, I use different event windows to test if this has impact in the results. Finally, I use aggregated ES score as independent variable.

7.1 Own regressions for each pillars

So far I have included all three pillars in the same regression. Now, I will include only one of the pillars in the differences-in-differences regression. In table 8 the first column has the regression where only the Environmental score is included in the regression. The second column has the results of the regression where only the Social score is involved, the third column has only the Governance score. The dependent variable is abnormal return.

The first column of table 8 shows that now the Environmental score has significant positive coefficient during the crisis at 5% level. These results suggests that the impact of the Environmental score is explained by the Social score, since in the regressions in which include all the pillars the Environmental score is not significant. In column two, the Social score is positive and significant with 1% level during crisis and also negative and statistically significant after the crisis. So, the Environmental and Governance scores explain away this after-crisis underperformance of firms with high Social score. Finally in the third column, the Governance score still has negative impact on returns during crisis and no significant impact after the crisis.

Table 6 *Own regressions for each pillar*

This table presents the results of a differences-in-differences regression estimating the impact of the ESG pillar scores on the stock returns during and after the COVID-19 market crash. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The sample period is from August 2019 to August 2020, starting before the crash and ending after the crash. Daily data is used. The model is:

$$Return_{i,t} = \beta_0 + \beta_1 ESG_{i,2019} * Crash_t + \beta_2 ESG_{i,2019} * Post\text{-}Crash_t + Time\ FE_t + Firm\ FE_i + \varepsilon_{i,t}$$

The dependent variable is the abnormal return (based on the Four Factor model) measured on monthly basis. $ESG_{i,2019}$ is either the Environmental, Social or the Governance dummy variable equal to one if the firm's score is in the top quartile. The ESG data is from Refinitv's database, measured at the end of 2019. $Crash_t$ is dummy equal to one during the days between February 24 and March 20. $Post\text{-}Crash_t$ is dummy set to one during the time period after the crisis from March 21 to the end of August. Control variables include factor loadings, the firm fixed effects, and time fixed effects on daily basis. All data is winsorized at the 1st and 99th percentiles. The Heteroskedasticity-consistent standard errors are presented in parentheses. The standard errors are clustered at firm level.

* p < 0.1, ** p < 0.05, *** p < 0.01

	Abnormal return (1)	Abnormal return (2)	Abnormal return (3)
*ENV * Crash*	0.13269**		
	(0.06676)		
*ENV * Post-Crash*	-0.02530		
	(0.02167)		
*SOC * Crash*		0.26795***	
		(0.06805)	
*SOC * Post-Crash*		-0.04443**	
		(0.02244)	
*GOV * Crash*			-0.17386**
			(0.07553)
*GOV * Post-Crash*			-0.00325
			(0.02465)
Time FE (daily)	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
SE clustered by	Firm	Firm	Firm
N	287040	287040	287040
Adj. R2	0.435	0.435	0.435

7.2 Redefining the ESG variables

In the above differences-in-differences regressions I have defined the firm as high ENV, SOC or GOV if the firm is in the top quartile. Now, I will lower this cut-off point by including all firms in the high responsibility category if the firm has higher score than the median. Thus, the firm will have ENV, SOC and GOV dummy variable set to one if the firm's respective score is higher than the median score of all the firms.

As the results show the social score has lost its statistical significance. This suggests that only very high Social score has the ability to protect firm's stock price during the market crash. It's not enough that the Social score is better than the median it must be in the top quartile. On the other hand, the coefficient for the Governance score has been able to keep its statistical significance, and as in the results above the impact is negative. And again, the Environmental score is not significant, but, interestingly, even though it is not significant at even 10% level the significance has increased compared to the regressions where the stricter definition of high-ESG scores were used. The post-crash coefficients are insignificant all around.

Table 7 *Differences-in-differences regression with different ESG variables*

This table presents the results of a differences-in-differences regression estimating the impact of the ESG pillar scores on the stock returns during and after the COVID-19 market crash. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The sample period is from August 2019 to August 2020, starting before the crash and ending after the crash. Daily data is used. The model is:

$$Return_{i,t} = \beta_0 + (\beta_1 ENV_{i,2019} + \beta_2 SOC_{i,2019} + \beta_3 GOV_{i,2019}) * Crash_t + (\beta_4 ENV_{i,2019} + \beta_5 SOC_{i,2019} + \beta_6 GOV_{i,2019}) * Post\text{-}Crash_t + Time\ FE_t + Firm\ FE_i + \varepsilon_{i,t}$$

The dependent variable is either raw return or abnormal return (based on the Four Factor model) measured on monthly basis. $ENV_{i,2019}$, $SOC_{i,2019}$ and $GOV_{i,2019}$ are the Environmental, Social and Governance dummy variables equal to one if the firm's score is larger than the median. The ESG data is from Refinitv's database, measured at the end of 2019. $Crash_t$ is dummy equal to one during the days between February 24 and March 20. $Post\text{-}Crash_t$ is dummy set to one during the time period after the crisis from March 21 to the end of August. Control variables include the firm fixed effects, and time fixed effects on daily basis. All data is winsorized at the 1st and 99th percentiles. The Heteroskedasticity-consistent standard errors are presented in parentheses. The standard errors are clustered at firm level.
* p < 0.1, ** p < 0.05, *** p < 0.01

	Raw return *(1)*	Abnormal return (2)
*ENV * Crash*	0.10491	0.10471
	(0.08434)	(0.08453)
*SOC * Crash*	0.13263	0.13741
	(0.08327)	(0.08348)
*GOV * Crash*	-0.15178**	-0.15549**
	(0.06777)	(0.06796)
*ENV * Post-Crash*	-0.03865	-0.03912
	(0.02571)	(0.02680)
SOC Post-Crash*	0.00562	0.01636
	(0.02548)	(0.02657)
*GOV * Post-Crash*	0.00210	-0.00625
	(0.02165)	(0.02260)
Time FE (daily)	Yes	Yes
Firm FE	Yes	Yes
SE clustered by	Firm	Firm
N	287040	287040
Adj. R2	0.250	0.435

7.3 Changing the event window

As the result so far point to the fact that there seems to be positive impact on the returns during the crash but the evidence for the hypothesized negative impact after the crisis is not evident, I'm going to change the post-crash event window. So far, the post-crash period has been defined as starting from March 21 and ending at the end of sample at the end of August 2020. But as can be seen from figure 1 the stock market recovery after the crisis stabilized during the summer. So, I'm going to determine new post-crash period starting from March 20 and ending the period at the end of June, shortening the period by two months to see if it affects the found impact during the recovery period. This is the same post-crash period that is used by Jurvanen (2020).

Table 10 provides the result of the regression, where the *Recovery* variable is dummy that equals one during the days from March 21 to the June 30 of 2020. Shortening the event window of the post-crash dummy variable to better reflect the recovery period does not affect the impact of the Social score during the recovery period. The coefficient is negative but not statistically significant. Thus, according to this analysis it doesn't seem likely that an investor in socially responsible firms has to pay for the "insurance" of better returns during the crash by poorer performance after the crash. Interestingly, now the coefficient for the Governance score is positive and statistically significant at 10% level for the regression with raw returns as dependent variable.

Table 8 *Differences-in-differences with different post-crash period*

This table presents the results of a differences-in-differences regression estimating the impact of the ESG pillar score on the stock returns during and after the COVID-19 market crash. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The sample period is from August 2019 to August 2020, starting before the crash and ending after the crash. Daily data is used. The model is:

$$Return_{i,t} = \beta_0 + (\beta_1 ENV_{i,2019} + \beta_2 SOC_{i,2019} + \beta_3 GOV_{i,2019}) * Crash_t + (\beta_4 ENV_{i,2019} + \beta_5 SOC_{i,2019} + \beta_6 GOV_{i,2019}) * Recovery_t + Time\ FE_t + Firm\ FE_i + \varepsilon_{i,t}$$

The dependent variable is either raw return or abnormal return (based on the Four Factor model) measured on monthly basis. $ENV_{i,2019}$, $SOC_{i,2019}$ and $GOV_{i,2019}$ are the Environmental, Social and Governance dummy variables equal to one if the firm's score is in the top quartile. The ESG data is from Refinitv's database, measured at the end of 2019. $Crash_t$ is dummy equal to one during the days between February 24 and March 20. $Recovery_t$ is dummy set to one during the time period after the crisis from March 21 to the end of June. Control variables include the firm fixed effects, and time fixed effects on daily basis. All data is winsorized at the 1st and 99th percentiles. The Heteroskedasticity-consistent standard errors are presented in parentheses. The standard errors are clustered at firm level.
* p < 0.1, ** p < 0.05, *** p < 0.01

	Raw return *(1)*	Abnormal return (2)
*ENV * Crash*	0.00332	0.00873
	(0.08303)	(0.08332)
*SOC * Crash*	0.34708***	0.34457***
	(0.08444)	(0.08471)
*GOV * Crash*	-0.25576***	-0.25934***
	(0.08058)	(0.08084)
*ENV * Post-Crash*	-0.02782	-0.00854
	(0.03428)	(0.03660)
SOC Post-Crash*	-0.03840	-0.04732
	(0.03599)	(0.03819)
*GOV * Post-Crash*	0.06054*	0.04779
	(0.03469)	(0.03638)
Time FE (daily)	Yes	Yes
Firm FE	Yes	Yes
SE clustered by	Firm	Firm
N	287040	287040
Adj. R2	0.251	0.435

7.4 Aggregated ES score

So far I have included the variables individually in the regression, but many of the past papers have used some aggregated measure of firm's social responsibility. Many of them have used only the Environmental and Social scores, not involving the Governance score in the measure. See for example Lins et al. (2016) and Albuquerque et al. (2020). The reason for leaving the Governance score outside of the measure has been that it is not really acknowladged as being part of firm's social responsibility. Thus, I have done robustness test by creating ES score. The *ES* score is calculated by giving both the Environmental and Social scores equal weights of 0.5. This weighting method was also used by Albuquerque et al. (2020).

As table 12 show the statistical significane is preserved when the two scores are combined and the Governance score is left out of the regression.

Table 9 *Differences-in-differences regression with ES score*

This table presents the results of a differences-in-differences regression estimating the impact of the ES score on the stock returns before, during and after the COVID-19 market crash. The sample includes all firms that are headquartered in Europe have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The sample period is from August 2019 to August 2020, starting before the crash and ending after the crash. Daily data is used. The model is:

$$Return_{i,t} = \beta_0 + \beta_1 ES_{i,2019} * Crash_t + \beta_2 ES_{i,2019} * Post\text{-}Crash_t + Time\ FE_t + Firm\ FE_i + \varepsilon_{i,t}$$

The dependent variable is either daily raw return or daily abnormal return (based on the Four Factor model). $ES_{i,2019}$ dummy variables equal to one if the firm's ES score is in the top quartile. The ES score is the average of the Environmental score and the Social score of the firm. The ES score is measured at the end of 2019 with data from Refinitiv. $Crash_t$ is dummy equal to one during the crash period from February 24 to March 20 and $Post\text{-}Crash_t$ is dummy set to one during the time period after the crisis. Control variables include the firm fixed effects, and time fixed effects on daily basis. All data is winsorized at the 1st and 99th percentiles. The Heteroskedasticity-consistent standard errors are presented in parentheses. The standard errors are clustered at firm level.
* p < 0.1, ** p < 0.05, *** p < 0.01

	Raw return *(1)*	Abnormal return (2)
*ES * Crash*	0.24008***	0.24077***
	(0.06524)	(0.06547)
*ES * Post-Crash*	-0.02451	-0.02296
	(0.02036)	(0.02154)
Time FE (daily)	Yes	Yes
Firm FE	Yes	Yes
SE clustered by	Firm	Firm
N	287040	287040
Adj. R2	0.250	0.435

8 Discussion of results

Results above suggests that only the Social score had positive impact on the returns during the COVID-19 crash in European markets. For the Environmental score I did not find any statistically significant effect. Past research has mostly focused on using aggregate measures of the Environmental and Social scores (e.g. Lins et al. (2016) and Albuquerque et al. (2020)) basically assuming, since these two scores are highly correlated, that differentiating these does not bring any added value to the study. And as it happens, both Lins et al. (2016) and Albuquerque et al. (2020) did not find different impacts between the two scores. Similarly, Ding et. al (2020) studying the COVID-19 crisis with global data set finds positive and significant impact for both social and environmental scores. Also, Demers et. al (2021) does not find difference between the two pillars. Interestingly, Broadstock et al. (2021) find that in Chinese stock market during the COVID-19 crisis the Environmental score and the Governance score impacted positively the returns, whereas the Social score had negative impact. Pastor and Vorsatz (2020) finds just the oppostive effect, compared to my findings, studying US mutual funds during COVID-19 crisis, namely that the Environmental score is the driving factor for outperformance. Also, Koskinen et al. (2019) studying the impact of ES in European markets during the 2008-2009 financial crisis finds that the Environmental score has significant and positive impact whereas the Social score has no impact. My results suggest that in Europe during the COVID-19 crash there was clear difference between the impact of the Environmental and Social scores, and the Social score was the driving factor for the outperformance of socially responsible firms. Comparing these results to the past findings mentioned above this result is unique.

Another interesting finding of mine was that the Governance score had statistically significant negative impact on the returns. Good governance has been found to be positively related to returns during the 2008 financial crisis by Lins et al. (2013) and Nguyen (2015). On the other hand, the results of Lins et al. (2016) suggest that the Governance score had no statistically significant impact on the returns during the financial crisis. Similarly Koskinen (2019) finds no significant impact of the Governance score during the financial crisis using

data comprised of European firms. Studying high sustainability equity mutual funds in US during the COVID-19 crisis Pastor and Vorsatz (2020) find no effect of Governance score on stock performance. On the other hand, studying the impact of ESG policies on returns gloablly during the COVID-19 crisis Ding et. al (2020) find negative impact for an "antitakeover devices" measure, where the antitakeover devices is part of the Governance score. In terms of these presented past findings, that mostly suggest either positive or neutral effect, the fact that my results suggest that good governance of the firm actually predicts underperformance during market crash suggest that there is ambiguity in the real impact of the firm's governance policies for stock returns during market downturns.

I also set out to test if the COVID-19 situation of the counrty, where the firm is located impacts the results. My findings provide evidence for the hypothesis that firms located in countries that were hit harder by COVID-19 pandemic during February and March 2020 benefitted from higher Social score than firms in other European countries. This could be because the firms located in these high COVID-19 countires experienced larger exogenous shock than other countries. Thus suggesting that the bigger the crash is the more meaningful the high Social score is for the resiliancy of firms' stock price.

One of the other hypothesis that I studied is the impact of ESG scores after the crash. Jurvanen (2020), and Nofsinger and Varma (2013) find that after the market downturn during the recovery period the environmentally and socially responsible stocks underperform other stocks. I did not find clear evidence for this effect. Only firms located in countries with worse COVID-19 situation and high Social score had negative impact after the crisis for having. This suggest that this post-crisis effect is only evident in a bigger crash.

There are number of technical limitations that affect this empirical research. There is the typical concern of endogeneity. The endogeneity problem is, at least partly, alleviated by the fact that the COVID-19 shock is exogenous since its source is outside of economic or financial concerns. Omitted variable bias is another problem. I have done my best to avoid the omitted variable bias by including control variables that have been shown to have significant explanatory power.

Another significant problem when studying the relationship between ESG and returns is the subjectivity in the ESG scores and the different methodologies in calculating the ESG score. See the literature review for further explanation on these problems. Since I use ESG scores by only one rating provider, it is possible that the results could differ if ratings provided by another rater were used. Albuquerque et al. (2020) find very similar results using MSCI ESG rates (previously known as KLD) when they compare them to the results discovered using the Refinitv ESG score.

It also must be noted that the importance of ESG is currently growing rapidly. This will change the relationship between ESG ratings and returns in the future. As more and more investors take ESG as important metric when making investment decisions all companies must take their ESG practices more seriously. When good ESG practices are taken as given by investors and companies, it's possible that the differences in ESG performances between companies is reduced and therefore it won't be as large factor in investment decisions. Also, as investors become more knowledgable of the positive effect of ESG on the firms profitability the market overperformance of high ESG stocks is likely to disappear.

All in all, comparing my findings to the findings of others leads me to conclude that the impact of ESG policies on returns during market crisis is at least partly ambigious. This thesis highlights two points related to this. First, when using the three pillars of ESG individually, rather than aggregating them into one measure, and comparing the individual effects, things are not as clear they first may have seem to be. Even though the correlation of these measures with each other is high, my results suggests there is significant difference how the three pillars impacted the returns. Second, my findings indicates that the effect is not geographically universal. My findings for the European firms do not correspond to the results found using US data in terms of the impact of each pillar separately. On the other hand, the impact can still be found when using the aggregate measure comprised of the Environmental and Social pillars suggesting that the same effect can be still be found but now it is driven by the Social pillar and not both pillars.

These findings advocate for the importance of differentiating between the different pillars of ESG rather than aggregating them to one measure. This also raises the question should the ESG be differentiated even further trying to find if there can be found the even more

nuanced differences in the impact, and thus helping the building of the theoretical foundations for the relationship between ESG and returns. Also, studying how this relationsip differs in different regions and trying to find if there is any theoretical explanations for these differences demands a wholistic research on this topic.

9 Conclusion

The COVID-19 pandemic caused extraordinary exogenous shock in the markets. Stock markets crashed all over the world when the seriousness of the virus began to become clear in investors' minds. But soon after the speedy collapse, when central banks and governments started to support economy and markets, the stock markets started to climb back up. In this thesis I use this exceptional time period to study the impact of ESG on stock returns in European markets during a market crash. Previous research has provided evidence for the fact that firms with better Environmental and Social performance has outperformed other firms in the stock market during market shocks (see Lins et. al (2016), Albuquerque et al. (2020) and Jurvanen (2020)). I find supporting evidence for these findings. Only difference is that I find positive and statisically significant parameter only for the Social score, indicating that firms with high Social performance outperformed other firms during the COVID-19 market crash. The Environmental score has no impact on the returns. The third pillar of ESG, Governance, had negative impact on returns. I also find that firms located in countries with larger exogenous shock, proxied by the number of reported COVID-19 cases, benefitted from the high Social score, whereas firms located in countries with less serious COVID-19 sitation there is no significant effect to be found. In addition, I find that after the crisis during the recovery period the firms with high social performance underperform other firms, but this is only evident for the firms located in countries with worse COVID-19 situation. These findings suggest that risk-based explanation for the effect is possible. By buying stocks' of firms with high Social score investor could protect himself better from large market crashes, but he has to pay for this outperformance by underperformaning after the crash.

This thesis has focused solely on the impact of ESG activities on the returns during the market crash leaving the possible explanations of this outperformance outside the scope of this study. As mentioned in the hypothesis construction section, there are already multiple hypotheses about the possible causes of the outperformance, e.g. ESG activities increasing social capital which makes firm performance more robust during market crisis (Lins et. al (2016), or product differentation as the result of ESG actitivities and thus making the firm perform better during market shocks. Testing these hypothesis could be an avenue for further research. Another possibility for further study could be to use different metric for the social responsibilty of a firm. As mentioned in the literature review there are large discrepancies in the ratings between different rating providers, and thus using different measure of social responsibilty could provide different results. In addition, the findings related to the differences in the impact of the three ESG pillar demands further research. Is this discrapency in the impact between the pillars explained by the dataset tied to certain geographical area and the time period studied? Or could these discrapencies be found more broadly? Is there any theoretical basis for these discrapencies?

References

Agudelo MAL, Jóhannsdóttir L, and Davídsdóttir B. 2019. "A literature review of the history and evolution of corporate social responsibility." *International Journal of Corporate Social Responsibility* 4:1.

Albuquerque R, Koskinen Y, and Zhang C. 2019. "Corporate social responsibility and firm risk: Theory and empirical evidence." *Management Science* 65:4451–69.

Albuquerque R, Koskinen Y, Yang S, and Zhang C. 2020. "Resiliency of Environmental and Social Stocks: An Analysis of the Exogenous COVID-19 Market Crash." *The Review of Corporate Finance Studies* 9(3): 593-621.

Barnea A, and Rubin A. 2005. "Corporate Social Responsibility as a Conflict between Owners." *Working Paper.*

Benson KL, and Humphrey JE. 2008. "Socially responsible investment funds: Investor reaction to current and past returns." *Journal of Banking & Finance* 32:1850-1859.

Berg F, Koelbel JF, Rigobon R. 2020. "Aggregate Confusion: The Divergence of ESG Ratings." *Working paper.*

Broadstock DC, Chan K, Cheng LTW, Wang X. 2021. "The role of ESG performance during times of financial crisis: Evidence from COVID-19 in China." *Finance Research Letters* 38:101716.

Cahart MM. 1997. "On Persistence in Mutual Fund Performance." *The Journal of Finance* 52(1):57-82.

Chatterji AK, Durand R, Levine DI, and Touboul S. 2016. "Do ratings of firms converge? Implications for managers, investors and strategy researchers." *Strategic Management Journal,* 37(8):1597–1614.

Cheng I, Hong H, and Shue K. 2013. "Do managers do good with other peoples' money?" *Working paper, Dartmouth College, Princeton University, and University of Chicago.*

Cheng IH, Hong H, and Shue K. 2013. "Do managers do good with other people's money?" *Working Paper.*

Clark GL, Feiner A, Viehs M. 2015. "From the stockholder to the stakeholder." *Working paper.*

Daniel K, Titman S. 1997. "Evidence on the characteristics of cross-sectional variation." *Journal of Finance* 52:1-33.

Demers E, Hendrikse J, Joos P, Lev B. 2021. "ESG Didn't Immunize Stocks During the COVID-19 Crisis,." *Journal of Business Finance & Accounting.*

Dhaliwal DS, Li OZ, Tsung A, and Yang YG. 2011. "Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting." *The Accounting Review* 86:1:59-100.

Ding W, Levine R, Lin C, Xie W. 2020. "Corporate immunity to the COVID-19 pandemic." *Working paper.*

Dooren BV, and Galema R. 2018. “Socially responsible investors and the disposition effect.” *Journal of Behavioral and Experimental Finance* 17:42-52.

Dorfleitner G, Halbritter G, and Nguyen M. 2015. “Measuring the level and risk of corporate responsibility – An empirical comparison of different ESG rating approaches.” *Journal of Asset Management* 16(7):450–466.

Döttling R, and Kim S. 2020. “Sustainability Preferences Under Stress: Evidence from Mutual Fund Flows During COVID-19.” *Working Paper.*

Eccles RG, Ioannou I, and Serafeim G. 2014. “The Impact of Corporate Sustainability on Organizational Processes and Performance.” *Management Science* 60:11:2835-2857.

Edmans, A. 2011. “Does the stock market fully value intangibles? Employee satisfaction and equity prices.” *Journal of Financial Economics* 101, 621–640.

Ferrell A, Liang H, and Renneboog L. 2016. “Socially responsible firms.” *Journal of Financial Economics* 122:585-606.

Folger-Laronde Z, Pashang S, Feor L, ElAlfy A. 2020. “ESG ratings and financial performance of exchange-traded funds during the COVID-19 pandemic.” *Journal of Sustainable Finance & Investment.*

Friede G, Busch T, and Bassen A. 2015. “ESG and financial performance: aggregated evidence from more than 2000 empirical studies.” *Journal of Sustainable Finance & Investment* 5:4: 210-233.

Friedman M. 1970. “The social responsibly of business is to increase its profits.” *New York Times Magazine* September 13.

Ghoul SG, Guedhami O, Kwok CCY, Mishra DR. 2011. “Does corporate social responsibility affect the cost of capital?” *Journal of Banking & Finance* 35:2388-2406.

Godfrey CG, Merrill CB, Hansen JM. 2009. “The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis.” *Strategic Management Journal* 30:425-445.

Harford J, Klasa S, Maxwell W. 2012. “Refinancing risk and cash holdings.” *Journal of Finance* 69:975-1012.

Hart O, Zingales L. 2017. “Companies Should Maximize Shareholder Welfare Not Market Value.” *Journal of Law, Finance, and Accounting* 2:247-274.

Hong H, Kubik JD, Liskovich I. 2019. “Crime, Punishment and the Value of Corporate Social Responsibility.” *Working paper.*

Jurvanen, H. 2020. “Corporate Social Responsibility as a Signal of Value During Market Shocks.” *Master's Thesis.*

Koskinen J. 2019. “Corporate Social Responsibilty and Stock Returns during the Financial Crisis: Evidene from Europe.” *Bachelor's thesis* Aalto University.

Krüger, P. 2015. “ Corporate goodness and shareholder wealth.” *Journal of Financial Economics* 115(2):304–329.

Li K, Liu X, Mai, F, Zhang T. 2021. "The Role of Corporate Culture in Bad Times: Evidence from the COVID-19 Pandemic." *Working Paper.*

Lins KV, Volpin P, and Wagner H. 2013. "Does family control matter? International evidence from the 2088–2009 financial crisis." *Review of Financial Studies* 26:2583-2619.

Lins, K. V., Servaes H., and Tamayo A. 2017. "Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis." *Journal of Finance* 72:1785–824.

Malik, M. 2015. "Value-Enhancing Capabilities of CSR: A Brief Review of Contemporary Literature." *Journal of Business Ethics* 127:419-438.

Margolis, J., Elfenbein A., and Walsh, P. 2009. "Does it pay to be good . . ..And Does it Matter? A Meta-Analysis of the Relationship between Corporate Social and Financial Performance." *Working paper, Harvard University, Washington University in St. Louis, and.*

Masulis RW, and Reza SW. 2015. "Agency problems of corporate philanthropy." *The Review of Financial Studies* 28(2):592–636.

McKinsey & Company. 2020. "The ESG premium: New Perspectives on value and performance." https://www.mckinsey.com/business-functions/sustainability/ourinsights/the-esg-premium-new-perspectives-on-value-and-performance.

Nguyen T, Nguyen HG, Yin X. 2015. "Corporate governance and corporate financing and investment during the 2007–2008 financial crisis." *Financial Management* 44:115-146.

Nofsinger J, Varma A. 2014. "Socially responsible funds and market crises." *Journal of Banking & Finance* 48:180-193.

P, Krüger. 2015. "Corporategoodnessandshareholderwealth." *Journal ofFinancialEconomics* 115:304-329.

Pagano M, Wagner C, Zechner J. 2020. "Disaster Resilience and Asset Prices." *Working Paper.*

Palma-Ruiz JM, Castillo-Apraiz J, Gomez-Martinez R. 2020. "Socially Responsible Investing as a Competitive Strategy for Trading Companies in Times of Upheaval Amid COVID-19: Evidence from Spain." *International Journal of Financial Studies* 8:41.

Pastor L, Vorsatz MB. 2020. "Mutual Fund Performance and Flows during the COVID-19 Crisis." *The Review of Asset Pricing Studies* 10:791-833.

PC, Godfrey. 2005. "The Relationship between Corporate Philanthropy and Shareholder Wealth: A Risk Management Perspective." *The Academy of Management Review* 30(4):777-798.

Ramelli S, Wagner AF. 2020. "Feverish stock price reactions to covid-19." *Working paper.*

Renneboog L, Horst JT, and Zhang C. 2011. "Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds." *Journal of Financial Intermediation* 20:562-588.

Renneboog L, Horst JT, and Zhang C. 2008. "The price of ethics and stakeholder governance: The performance of socially responsible mutual funds." *Journal of Corporate Finance* 14:302-322.

Semenova N, and Hassel LG. 2015. “On the Validity of Environmental Performance Metrics.” *Journal of Business Ethics* 132(2):249–258.

Servaes H, and Tamayo A. 2013. “The impact of corporate social responsibility on firm value: The role of customer awareness.” *Management Science* 59(5):1045–1061.

Singh A. 2020. “COVID-19 and safer investment bets.” *Finance Research Letter* 36:101729.

Takashi H, Yamada K. 2021. “When the Japanese stock market meets COVID-19: Impact of ownership, China and US exposure, and ESG channels.” *International Review of Financial Analysis* 101670.

Appendix

Table 10 *Cross sectional regression: Relation between ES and stock returns*

This table presents the results of a cross-sectional regression estimating the relation between ES and stock returns during the COVID-19 market crash. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The model is as follows:

$$Crash\ Return_i = \beta_0 + \beta_1 ES_i + \beta_2 Firm\ Controls_i + Industry\ FE_i + Country\ FE_i + \varepsilon_i.$$

The dependent variable is either raw return or abnormal return (based on the Four Factor model) from February 24 to March 20. The independent variables include the ES score and control variables. Columns 1 and 2 do include only the industry fixed effects and country fixed effects as control variables. The columns 3 and 4 add various firm characteristics as control variables (data from the end of 2019). These include Cash Holdings (cash and marketable securities divided by assets), Profitability (operating income divided by assets), Short-Term Debt (short-term debt divided by assets) and Long-Term Debt (long-term debt divided by assets), Size (log of firm's market value of equity), Book-to-Market (book value divided by market value of equity), Momentum (12-month return prior to the crash period), dummy for Negative Book-to-Market ratio and Idiosyncratic Risk (residual variance of the market model). The Heteroskedasticity-consistent standard errors are presented in parentheses. All data winsorized at the 1st and 99th percentiles. * p < 0.1, ** p < 0.05, *** p < 0.01

	Raw return	Abnormal return	Raw return	Abnormal return
	(1)	(2)	(3)	(4)
ES	0.022	0.021	0.013	0.079***
	(0.019)	(0.023)	(0.020)	(0.028)
Size			-0.002	-0.005
			(0.003)	(0.004)
Long-Term Debt			-0.038	-0.039
			(0.023)	(0.035)
Short-Term Debt			0.021	0.059
			(0.060)	(0.090)
Cash holdings			0.026	0.159***
			(0.038)	(0.058)
Profitability			-0.003	0.017
			(0.049)	(0.074)
Book-to-Market			0.003	0.073***
			(0.011)	(0.016)
Negative B/M			-0.026	0.035
			(0.032)	(0.053)
Momentum			0.359***	0.247***
			(0.017)	(0.023)
Idiosyncratic Risk			-2.342***	2.841**
			(0.713)	(1.234)
N	1104	1104	1104	1104
Adj. R^2	0.183	0.104	0.536	0.213
Country FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes

Table 11 *Differences-in-differences regression: Relation between linear ESG measures and stock returns*

This table presents the results of a differences-in-differences regression estimating the impact of the ESG pillar scores on the stock returns before, during and after the COVID-19 market crash. The sample includes all firms that are headquartered in Europe, have larger market value than 250M€, and all the necessary data are available on Datastream. Altogether 1104 firms are included in the sample. The sample period is from August 2019 to August 2020, starting before the crash and ending after the crash. Daily data is used. The model is:

$$Return_{i,t} = \beta_0 + (\beta_1 ENV_{i,2019} + \beta_2 SOC_{i,2019} + \beta_3 GOV_{i,2019}) * Crash_t + (\beta_4 ENV_{i,2019} + \beta_5 SOC_{i,2019} + \beta_6 GOV_{i,2019}) * Post\text{-}Crash_t + Time\ FE_t + Firm\ FE_i + \varepsilon_{i,t}$$

The dependent variable is either raw return or abnormal return (based on the Four Factor model) measured on monthly basis. $ENV_{i,2019}$, $SOC_{i,2019}$ and $GOV_{i,2019}$ are equal to the Environmental, Social or Governance score of the firm. The ESG data is from Refinitv's database, measured at the end of 2019. $Crash_t$ is dummy equal to one during the days between February 24 and March 20. $Post\text{-}Crash_t$ is dummy set to one during the time period after the crisis from March 21 to the end of August. Control variables include the firm fixed effects, and time fixed effects on daily basis. All data is winsorized at the 1st and 99th percentiles. The Heteroskedasticity-consistent standard errors are presented in parentheses. The standard errors are clustered at firm level.
* p < 0.1, ** p < 0.05, *** p < 0.01

	Raw return *(1)*	Abnormal return (2)
*ENV * Crash*	0.00039	0.00033
	(0.00196)	(0.00197)
*SOC * Crash*	0.00509**	0.00527**
	(0.00233)	(0.00234)
*GOV * Crash*	-0.00401**	-0.00412**
	(0.00172)	(0.00172)
*ENV * Post-Crash*	-0.00049	-0.00062
	(0.00059)	(0.00063)
SOC Post-Crash*	-0.00068	-0.00027
	(0.00073)	(0.00077)
*GOV * Post-Crash*	0.00079	0.00056
	(0.00052)	(0.00054)
Time FE (daily)	Yes	Yes
Firm FE	Yes	Yes
SE clustered by	Firm	Firm
N	287040	287040
Adj. R2	0.250	0.435